

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2005, 21:55:52 ; Search time 210.545 Seconds
(without alignments)
17070.053 Million cell updates/sec

Title: US-09-551-494-5_COPY_5430_5505

Perfect score: 76

Sequence: 1 ggcacagagctcgccaat.....tgaagtaaccaatgctgtga 76

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*
1: gb_ba.*
2: gb_htg.*
3: gb_in.*
4: gb_om.*
5: gb_ov.*
6: gb_pat.*
7: gb_ph.*
8: gb_pl.*
9: gb_pr.*
10: gb_ro.*
11: gb_sts.*
12: gb_ay.*
13: gb_un.*
14: gb_vi.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	76	100.0	6355	6 AX040177	AX040177 Sequence
2	76	100.0	6355	6 AX795380	AX795380 Sequence
3	76	100.0	6355	14 TMGCG	M34077 Tobacco mil
4	71.2	93.7	1413	14 TMGMPCPA	M34236 Tobacco mil
5	69.6	91.6	6356	14 AB078435	AB078435 Tobacco m
6	58.6	77.1	411	6 AX040184	AX040184 Sequence
7	45.8	60.3	912	14 ORVTRANS	X55296 Odontogloss
8	45.8	60.3	5997	6 E03624	E03624 DNA encodin
9	45.8	60.3	6597	6 E04305	E04305 cDNA encodi
10	45.8	60.3	6609	14 ORU34586	U34586 Odontogloss
11	45.8	60.3	6611	14 S83257	S83257 126 kda pro
12	45.8	60.3	6612	14 AV571290	AV571290 Odontoglo
13	44	57.9	1004	14 TOTWV2	V01406 TMV 3' end.
14	44	57.9	1019	6 E00088	E00088 Genomic rna
15	44	57.9	6395	14 D63809	D63809 Tobacco mos
16	42.4	55.8	396	14 TM0307579	AJ307579 Tobacco m
17	42.4	55.8	632	6 A68760	A68760 Sequence 1
18	42.4	55.8	706	14 TM0308689	AJ308689 Tobacco m
19	42.4	55.8	750	14 TM0509080	AJ509080 Tobacco m

20	42.4	55.8	750	14 TM0509084	AJ509084 Tobacco m
21	42.4	55.8	782	14 TM0308692	AJ308692 Tobacco m
22	42.4	55.8	788	14 TM0308685	AJ308685 Tobacco m
23	42.4	55.8	790	14 TM0307583	AJ307583 Tobacco m
24	42.4	55.8	804	6 CQ793025	CQ793025 Sequence
25	42.4	55.8	804	14 TM0307582	AJ307582 Tobacco m
26	42.4	55.8	806	14 TM0308682	AJ308682 Tobacco m
27	42.4	55.8	806	14 TM0308684	AJ308684 Tobacco m
28	42.4	55.8	807	6 BD263879	BD263879 Viral exp
29	42.4	55.8	807	6 BD263880	BD263880 Viral exp
30	42.4	55.8	807	6 AR435699	AR435699 Sequence
31	42.4	55.8	807	6 AR435700	AR435700 Sequence
32	42.4	55.8	807	6 AX045758	AX045758 Sequence
33	42.4	55.8	807	6 AX592974	AX592974 Sequence
34	42.4	55.8	807	14 TM0307581	AJ307581 Tobacco m
35	42.4	55.8	807	14 TM0308683	AJ308683 Tobacco m
36	42.4	55.8	807	14 TM0308693	AJ308693 Tobacco m
37	42.4	55.8	807	14 TM0509081	AJ509081 Tobacco m
38	42.4	55.8	808	14 TM0307578	AJ307578 Tobacco m
39	42.4	55.8	808	14 TM0308686	AJ308686 Tobacco m
40	42.4	55.8	808	14 TM0308688	AJ308688 Tobacco m
41	42.4	55.8	808	14 TM0308690	AJ308690 Tobacco m
42	42.4	55.8	809	14 TM0308691	AJ308691 Tobacco m
43	42.4	55.8	811	14 TM0509082	AJ509082 Tobacco m
44	42.4	55.8	815	14 AY300161	AY300161 Tobacco m
45	42.4	55.8	817	14 TM0509083	AJ509083 Tobacco m

ALIGNMENTS

RESULT 1	AX040177	Sequence 5 from Patent WO0063397.	6355 bp	DNA	linear	PAT 18-NOV-2000
LOCUS	AX040177	Sequence 5 from Patent WO0063397.				
DEFINITION	AX040177					
ACCESSION	AX040177					
VERSION	AX040177.1	GI:11230127				
KEYWORDS						
SOURCE		synthetic construct				
ORGANISM		synthetic construct				
REFERENCE	1	artificial sequences.				
AUTHORS		Meulewaeter, F., Cornelisse, Jacobs, J., van Eldik, G. and Metzlaiff, M.				
TITLE		Methods and means for delivering inhibitory rna to plants and appl				
JOURNAL		ications thereof				
FEATURES		Patent: WO 0063397-A 5 26-OCT-2000;				
		Aventis CropScience N.V. (BE)				
		Location/Qualifiers				
		1. .6355				
		/organism="synthetic construct"				
		/mol_type="unassigned DNA"				
		/db_xref="taxon:32630"				
		/notes="cDNA copy of the nucleotide sequence of the genome				
		of TMV-U2"				
ORIGIN						
		Query Match	100.0%;	Score 76;	DB 6;	Length 6355;
		Best Local Similarity	100.0%;	Pred. NO. 2.2e-15;		
		Matches	76;	Conservative	0;	Mismatches 0;
				Indels	0;	Gaps 0;
Qy	1	GTGACAGACGGCTCGCCAAATTTGAACCTGACCTGAAAGCTTTGTGAGGAGTTGCTGGATGAA	60			
Db	5430	GTGACAGACGGCTCGCCAAATTTGAACCTGACCTGAAAGCTTTGTGAGGAGTTGCTGGATGAA	5489			
Qy	61	GTACCAATGGCTGTGA	76			
Db	5490	GTACCAATGGCTGTGA	5505			
RESULT 2						
AX795380		Sequence 7 from Patent WO03052108.	6355 bp	mRNA	linear	PAT 04-OCT-2003
LOCUS						
DEFINITION						

/note="coat protein"	
/codon_start=1	
/protein_id="AAA47937.1"	
/db_xref="GI:335247"	
/translation="MPVTINSPSQFVYLSSAYADPVQLINLCTNALGNQFQTQOARTT	
VQQQFADAMKFPVSMTRVPASDFYVYRNSITLDPILITALLNSFDTRNRIIEVDNQPA	
PNTTEIWNATORVDDATVAIRASINNLANELVRGTGMFNQAFETASGLVWTTTTPAT"	
ORIGIN	
Query Match	100.0%; Score 76; DB 14; Length 6355;
Best Local Similarity	100.0%; Pred. No. 2.2e-15;
Matches	76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Oy	1 GTGACAGACGGCTGCGCAATGAACCTCACTGAAAAAGGTTGTTGAGGAGTTCTGGATGAA 60
Db	5430 GTGACAGACGGCTGCGCAATGAACCTCACTGAAAAAGGTTGTTGAGGAGTTCTGGATGAA 5489
Oy	61 GTACCAATGGCTGTGA 76
Db	5490 GTACCAATGGCTGTGA 5505
RESULT 4	
TMGMPCPA	1413 bp ss-RNA linear VRL 03-AUG-1993
LOCUS	Tobacco mild green mosaic virus movement and coat protein genes,
DEFINITION	complete cds.
ACCESSION	M34236.1 GI:335248
VERSION	coat protein; movement protein.
KEYWORDS	Tobacco mild green mosaic virus
SOURCE	Tobacco mild green mosaic virus
ORGANISM	Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE	1 (bases 1 to 1413)
AUTHORS	Nejidat,A., Cellier,F., Holt,C.A., Gafny,R., Eggenberger,A.L. and
TITLE	Beachy,R.N.
JOURNAL	Transfer of the movement protein gene between two tobamoviruses:
MEDLINE	influence on local lesion development
PUBMED	Virology 180 (1), 318-326 (1991)
COMMENT	91082424
FEATURES	1984654
source	Original source text: Tobacco mild green mosaic virus (strain PV
	228), cDNA to viral RNA.
	Draft entry and computer-readable sequence for [1] kindly submitted
	by A.Nejidat, 11-MAY-1990.
	Location/Qualifiers
	1. .1413
	/organism="Tobacco mild green mosaic virus"
	/mol_type="genomic RNA"
	/db_xref="taxon:12241"
	51. .851
CDS	/codon_start=1
	/product="movement protein"
	/protein_id="AAA47938.1"
	/db_xref="GI:335249"
	/translation="MAVSLRDTVKISFINLSKQDEILPAFMTKVKSVRISTVDKIMA
	VKXNSLSDVLKGLVKVNGYVCLAGLVVSGWNLDPNCRGVSVCI VDKMKRNGE
	ATIGAYHAPACKNFSPKLPNYSTSEDAEKPMOVLVNI KGVAMEEGYCPISLEFV
	SICVWHKNNVKGRLERILRVTDSDSPIELTEKVEEFPVDPVPAVKLERFRPKTKXGK
	RKKEKKRVVGVSNVNNKNNINNSGKKGLKVEIEDNVSDDESIIASSSTF"
	857. .1336
CDS	/codon_start=1
	/product="coat protein"
	/protein_id="AAA47939.1"
	/db_xref="GI:335250"
	/translation="MPVTINSPSQFVYLSSAYADPVQLINLCTNALGNQFQTQOARTT
	VQQQFADAMKFPVSMTRVPASDFYVYRNSITLDPILITALLNSFDTRNRIIEVDNQPA
	PNTTEIWNATORVDDATVAIRASINNLANELVRGTGMFNQAFETASGLVWTTTTPAT"
ORIGIN	
Query Match	93.7%; Score 71.2; DB 14; Length 1413;
Best Local Similarity	96.1%; Pred. No. 7.6e-14;
Matches	73; Conservative 0; Mismatches 3; Indels 0; Gaps 0;


```

mat_peptide 1..909
ORIGIN
  /product="cell-to-cell transport protein"

Query Match 60.3%; Score 45.8; DB 14; Length 912;
Best Local Similarity 76.7%; Pred. No. 4.8e-05;
Matches 56; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GTGACAGCGCTCGCCAAATGAACCTCACTGAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
Db 613 GTAACAGAGAGGCGCCACCGAAGCTTACTGAACAGTGTGTGATGATGTTCTGTGGAGAA 672

Qy 61 GTACCAATGGCTG 73
Db 673 GTTCCTATGGCTG 685

RESULT 8
E03624
LOCUS
DEFINITION DNA encoding a part of complementary DNA to odontoglossum ringspot
virus genomic RNA.
ACCESSION E03624
VERSION E03624.1 GI:2171839
KEYWORDS JP 1992144685-A/1.
SOURCE Odontoglossum ringspot virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE 1 (bases 1 to 5997)
AUTHORS Isomura,Y., Matsumoto,Y., Chatani,M. and Ikegami,M.
TITLES CDNA OF ORSV GENE
JOURNAL NIPPON OIL CO LTD
COMMENT Patent: JP 1992144685-A 1 19-MAY-1992;
OS odontoglossum ringspot virus
PN JP 1992144685-A/1
PD 19-MAY-1992
PF 28-NOV-1989 JP 1989306626
PI IKEGAMI MASATO
PC C12N15/40.C07K13/00.C07K15/04.C12N1/21.C12N5/10.C12P21/02. PC
(C12P21/02.
PC C12R1.19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pORES-25;
FH Key Location/Qualifiers
FT misc_feature 1..5997
FT /note='a part of complementary DNA to FT
FT Odontoglossum
FT ringspot virus genomic RNA'
FT /note='a part of complementary DNA to ORSV'.
FEATURES
source
  Location/Qualifiers
  1..5997
  /organism="Odontoglossum ringspot virus"
  /mol_type="genomic RNA"
  /db_xref="taxon:12238"

ORIGIN
  /product="cell-to-cell transport protein"

Query Match 60.3%; Score 45.8; DB 6; Length 5997;
Best Local Similarity 76.7%; Pred. No. 7.1e-05;
Matches 56; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GTGACAGCGCTCGCCAAATGAACCTCACTGAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
Db 4805 GTAACAGAGAGGCGCCACCGAAGCTTACTGAACAGTGTGTGATGATGTTCTGTGGAGAA 4864

Qy 61 GTACCAATGGCTG 73
Db 4865 GTTCCTATGGCTG 4877

RESULT 9
E04305
LOCUS
DEFINITION CDNA encoding Odontoglossum ring spot virus (ORSV) genomic RNA.
ACCESSION E04305
VERSION E04305.1 GI:2172508
KEYWORDS JP 1993030975-A/1.
SOURCE Odontoglossum ringspot virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE 1 (bases 1 to 6597)
AUTHORS Isomura,Y., Matsumoto,Y., Chatani,M., Mizuta,Y. and Ikegami,M.
TITLES CDNA OF ORSV GENE
JOURNAL NIPPON OIL CO LTD
COMMENT Patent: JP 1993030975-A 1 09-FEB-1993;
OS Odontoglossum ring spot virus
PN JP 1993030975-A/1
PD 09-FEB-1993
PF 26-JUL-1991 JP 1991276075
PI ISOMURA YOSHIKATSU, MATSUMOTO YOSHITOMO, CHATANI MASAOKI, PI
MIZUTA YOSHINORI, IKEGAMI MASATO
PC C12N15/40.C07K15/04.C12N1/21.C12N15/11.C12N15/70.C12P21/02. PC
C12Q1/68.
PC (C12N1/21.C12R1.19). (C12N15/70.C12R1.19). (C12P21/02.C12R1.19);
CC strandedness: Double;
CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
CC *source: clone=pORCP-D83;
FH Key Location/Qualifiers
FT misc_feature 1..6597
FT /note='cdna to Odontoglossum ring spot virus
FT genomic RNA'
FT 5707..6183
FT CDS
FT /product='coat protein'
FT mat_peptide 5707..6180
FT /product='coat protein'.
FEATURES
source
  Location/Qualifiers
  1..6597
  /organism="Odontoglossum ringspot virus"
  /mol_type="genomic RNA"
  /db_xref="taxon:12238"

ORIGIN
  /product="cell-to-cell transport protein"

Query Match 60.3%; Score 45.8; DB 6; Length 6597;
Best Local Similarity 76.7%; Pred. No. 7.3e-05;
Matches 56; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GTGACAGCGCTCGCCAAATGAACCTCACTGAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
Db 5405 GTAACAGAGAGGCGCCACCGAAGCTTACTGAACAGTGTGTGATGATGTTCTGTGGAGAA 5464

Qy 61 GTACCAATGGCTG 73
Db 5465 GTTCCTATGGCTG 5477

RESULT 10
ORU34586
LOCUS
DEFINITION Odontoglossum ringspot virus, complete genome.
ACCESSION ORU34586
VERSION U34586.1 GI:1407591
KEYWORDS Odontoglossum ringspot virus
SOURCE Odontoglossum ringspot virus
ORGANISM Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE 1 (bases 1 to 6609)
AUTHORS Chng,C.G., Wong,S.M., Mahtani,P.H., Loh,C.S., Goh,C.J., Kao,M.C.,
Chung,M.C. and Watanabe,Y.
TITLES The complete sequence of a Singapore isolate of odontoglossum

```


SAGLEENLAVAMIKRNFNSPDLTGILIEDTAELVNVKFWDAYIIIDELSGGNVTMTSD
AFHRMAQKESITIGQLADFDVFLPAIDQYKMIKAQPKQKLDLSPODEYAAQTV
YHSQINAFIQLRQLERIDSSKFLFYTRKPEQIEBFFSOLDSTVPMVEVLE
LDISKYDSQNEFHCAVYLIIWEKLGFLFEEVWKQGRKTSIKTAGTKICLWQ
RKSQDVITIGNTVIAACLASMIKPMKVIKAAFCGDDSIYIPKGLDLPDIQSGANL
MNFPAKLYRKYGVFCGRYIIHHRGAIIVYDPLKLSKLGCKHLSLDHLEFRIS
LCDYSSSLANNCAYFQGLNDIAEYHKTAVNGSPAFCSIVKYSKLNLFRTILFYNGSST
KG"

CDS

64. .3402
/gene="RdRp"
/codon_start=1
/product="RNA-dependent RNA polymerase/replicase"
/protein_id="AAS87223.1"
/db_xref="GI:46309860"
/translation="MAHFQQTMMNNKIVIEAGMRNSLINDLAQRVYDYNVAEELNHRSR
RPKNYSKVIQSEQIIQATNAYPEFEITFYNTQAVHSMAGGLRLELEYLMQIIPFG
SVTYDIGNFSAHLYKGRDYVHCWPNLDIRDVARHINQOQTVSTYTLARLERSKGLP
VFQOSAFNKYANDPDVCCDKRFODCSYVDLPCKTYAVALHSYDIIPADEFGAALLR
KDVHICVAAHFSENLLETSSAELDEIGATFYKSGRLSFFQNESTLNTYHSYKNV
IKYVCKTFFPASNPVYHKEFMCRTVNTFCFKTKVDYTYFLFRGVYTRGDSSEQFYTA
MDEAWYKKTLLAMNSERTIPDRRAAVNFNFPKDMVIIVLPFGSVTSGMKRSEYV
NKDFVYTLVNHRTYODKALTYKNVLSFVESIRSVIINGVTARSEWDVDSVLQAL
SMTEFLOTKLAADQVVLKFKFDDTVNLFWKOISDAVDLFPSPKRLISGGFV
KVARSQIQTPTDEVITFADKLVMEYKATELOHLDISKPLERAKEYNALSELVYK
ECDFEDITQFNLCDEKIDDPVAVKVIIVAKMNELILPKNPTPEALSALSPLPKD
LDMFDDLLKSTCAFPFVSFKTLDSGLLPKQSYGDERQFESQSVSVSDPHLKSYESVK
MKMWSAVYTGPKVQKQMNMYDLSASISATVSNLCKVLKVDYTGADPESEKSGYVD
VVGKLLPKDKCHAWGVAELNGEKVIIVLEWADGFPICGDMWRVAVSDSLIYSD
MKQTLTLLSCLKDGCEPVPSEDAKVTLDVPGCGCKTKBILFTVNFDEDLILVPGKEACK
MIIKRANKSGHVRATKINRVTVDSFLMHLKPKTKYKLFIDEGMLHTGCVNFIALSH
CREAMVGDTEQIIPKINRVANFPYKHPATLYDHREVRRLSLKCPADVTFHNSKYD
GKVLCTSDVIAEVRVGVNGVNPSPKPLKGIITFTQSDKAELKHEVEHVSFTG
EVNTVHEIQGTFFEDVSVRITPTPLELISKSSPHVLVALTRHTKSPKYYSVVLDPLV
KVCSDLKSVSDFLDMYKVDAGIL"

gene

CDS

4808. .5719
/gene="Wp"
/codon_start=1
/product="movement protein"
/protein_id="AAS87225.1"
/db_xref="GI:46309862"
/translation="WGLRVRVLLSIPIKTFEPSPCTMALVLDSDIKISEFINLSAS
EKLPSALTAVKSVIRISKVDKIISYENDTLDIDLLAGVKLVENGYVCLAGLVYTGW
NLPCNKGVSICLVDRKMRANEATIGSYHTSACKRFTFKIPNYSVTVDALGKI
WQVNTNIRGVEMEKFCPLSLEFVSIIVVNNIKLGRKILNVTGGTELTAVV
DEFVKVPMARLKSFRSNNKSPNSKFNKGRSLNRSNKLNYENGSDVGTSVVD
DIVVSGVSDIRIDDDCESFDAQSDSY"
5722. .6198
/gene="Cp"
/codon_start=1
/product="coat protein"
/protein_id="AAS87226.1"
/db_xref="GI:46309863"
/translation="MSYTIIDPSKAYLSSAWADPNLSLINLTSLNGLNQFQTOQARTV
VQOQADVWQFVPTLTSRFPAGQYFVRYDPLDPLITFLMGTFDTRNRIIEVENQAN
QNPTTTETLDRVDDATVAIRSAINNLLNELVGRGTMYNQVSEFETMSGLTWTSS"

gene

CDS

5722. .6198
/gene="Cp"
/codon_start=1
/product="coat protein"
/protein_id="AAS87226.1"
/db_xref="GI:46309863"
/translation="MSYTIIDPSKAYLSSAWADPNLSLINLTSLNGLNQFQTOQARTV
VQOQADVWQFVPTLTSRFPAGQYFVRYDPLDPLITFLMGTFDTRNRIIEVENQAN
QNPTTTETLDRVDDATVAIRSAINNLLNELVGRGTMYNQVSEFETMSGLTWTSS"

ORIGIN

Query Match 60.3%; Score 45.8; DB 14; Length 6612;
Best Local Similarity 76.7%; Pred. No. 7.3e-05;
Matches 56; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GTGACAGCGCTCGCCAAATTGAACCTCACTGAAAGGTTGTTGAGGAGTTCTGTTGAGTAA 60
Db 5420 GTACAGAGAGAGGCGCCACCAACTTACTGACGACGTTGTTGATGAGTTCTGTTGAGTAA 5479
Qy 61 GTACCAATGGCTG 73
Db 5480 GTTCCTATGCTG 5492

RESULT 13

TOTW2

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

COMMENT

FEATURES

source

1. .1004

/organism="Tobacco mosaic virus"

/mol_type="genomic RNA"

/db_xref="taxon:12242"

321. .800

/note="unnamed protein product; reading frame (coat protein)"

/codon_start=1

/protein_id="CAA24695.1"

/db_xref="GI:62120"

/db_xref="GOA:P03570"

/translation="MSYSITTPSQFVFLSSAWADPIELINLTNALGNQFQTOQARTV
VQOQSEVWKPSPQTVRPPDSDFKVRVNAVLDPLVTLGAPDTRNRIIEVENQAN
PTTAETLDRVDDATVAIRSAINNLLNELVGRGTMYNQVSEFESSGLVWTSPAT"

ORIGIN

Query Match 57.9%; Score 44; DB 14; Length 1004;

Best Local Similarity 73.7%; Pred. No. 0.00021;

Matches 56; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 GTGACAGCGCTCGCCAAATTGAACCTCACTGAAAGGTTGTTGAGGAGTTCTGTTGAGTAA 60

Db 52 GTGACAGAGCGAGCGGCCCATGGAACCTACAGAGAAAGTTGTTGATGAGTTCTATGGAAGAT 111

Qy 61 GTACCAATGGCTGCTGA 76

Db 112 GTCCCTATGTCATCA 127

RESULT 14

E00088

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

COMMENT

1 (bases 1 to 1019)

Roorensu, I.P. and Merii, K.H.

RNA PLANT VIRUS VECTOR OR FABRICATION AND UTILIZATION OF PART

Patent: JP 1983051894-A 2 26-MAR-1983;

NATL RES KAUNSURU OBU CANADA

OS tobacco mosaic virus

PN JP 1983051894-A/2

PD 26-MAR-1983

PF 27-MAY-1982 JP 1982090482

PI ROORENSU II PERUCHIYAA, MERII KURISUTEIN HARASA PC

CI2N15/00.A01H1/00.C07H21/04.C12P19/34.C12P21/00//CI2R1/91; CC

strandedness: Single;

```

CC topology: Linear;
CC hypothetical: No;
CC anti-sense: No;
FH Key
FT 5'UTR 1..320
FT misc_feature 61..234
FT nuclear' region essential for formation of FT
FT misc_feature 235..320
FT /note='region for regulation' FT CDS
FT 321..749
FT /product='capsid protein'
FT 3'UTR 750..1022.
FT Location/Qualifiers
FT 1..1019
FT /organism="Tobacco mosaic virus"
FT /mol_type="genomic RNA"
FT /db_xref="taxon:12242"

FEATURES
source
Query Match 57.9%; Score 44; DB 6; Length 1019;
Best Local Similarity 73.7%; Pred. No. 0.00021;
Matches 56; Conservative 0; Mismatches 20; Indels 0; Gaps 0;

Qy 1 GTGACAGCGCTCGCAATGTGAACCTCACTGAAAGCTTGTGAGGAGTTCGTGGATGAA 60
Db 52 GTGAGACGCGAGCGCCCTGGAACCTTACAGAGAGTGTGTGATGAGTTCATGGAAGAT 111
Qy 61 GTACCAATGGCTGTGA 76
Db 112 GTCCCTATGTCAATCA 127

RESULT 15
D63809
LOCUS D63809 6395 bp RNA linear VRL 13-FEB-1999
DEFINITION Tobacco mosaic virus genomic RNA for 130K protein, 180K protein,
30K protein and coat protein, complete sequence.
ACCESSION D63809.1 GI:1619995
VERSION coat protein; 30K protein; 180K protein; 130K protein.
KEYWORDS tobacco mosaic virus
SOURCE tobacco mosaic virus
ORGANISM
VIRUSES; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
Chen, J., Watanabe, Y., Sako, N., Ohshima, K. and Okada, Y.
Complete nucleotide sequence and synthesis of infectious in vitro
transcripts from a full-length cDNA clone of a rakkyo strain of
tobacco mosaic virus
JOURNAL Arch. Virol. 141 (5), 885-900 (1996)
MEDLINE 96265021
PUBMED 8678834
REFERENCE 2 (bases 1 to 6395)
AUTHORS Sako, N.
DIRECT SUBMISSION
TITLE Submitted (07-AUG-1995) Nobumichi Sako, Faculty of Agriculture,
Saga University, Laboratory of Plant Virology; 1 Honjo-machi, Saga,
Saka 840, Japan (Tel:0952-24-5191(ex.2730), Fax:0952-22-6274)
LOCATION/Qualifiers
1..6395
/organism="Tobacco mosaic virus"
/mol_type="genomic RNA"
/strain="Rakkyo"
/db_xref="taxon:12242"
69..4919
/note="position 3417..3419 is readthrough codon"
/codon_start=1
/transl_except=(pos:3417..3419,aa:OTHER)
/product="180K protein"
/protein_id="BAA09876.1"
/db_xref="GI:1619996"
/translation="MAYTQTATTSALLDVTVRGNNSLVNDLAKRLRYDTAVDEFNARDR"

FEATURES
source
CDS

```

```

RPKVNFSKVISBEBQTLIATRAYPEFQITFYNTQNAVHSLAGLSRSLSELYLMMQIPYG
SLTYDIGNPASHLFGKRAYVHCCMPNLDVRDLMRHEGQKDSIELYSLRDRGKGTVP
NFOKEAPDRYAEPEPDAVCHNTQTCEHQPQQSQKYYAIALHSIVDIPADEFGAALL
RKNVTCYAAPHFSENLLEDSSVNLDEINACFSRDGDKLTFFSASSTLYNCHFSFN
ILKYVCKTYPPASNRREYVYMKFVLVNTVTFCKFSRIDTFLYKGVAAHKSVDSSQFYT
AMEDAWHYKKTLCAMCNSERILLEDDSSVNTWFFPKRDMVIIPLPDI SLETSKRKEV
LVSKDQFVTLNHIPTYQAKALTYANLVSEFVSIRSRVINGVTARSEMDVDKSLQS
LSMTFFLHTKLAVLKDDLLISKFSGLSKTKVCOHWDEISLAFGNAPPSVKERLANKL
IRAAGDALEIKVPDLVITFHDRLVAEYKSSVDMPALDIRKMEETEVMYNALSLSL
RESKDFDVFSQCKSLEVDPMTAAKVIIVAVMSNESGLTILTPERPEANVAQAALDQ
EKASGALVVTSSREVPEEPKQSMGELQLAGFAGHPSSYRNKEIEISLEGFHMA
TADSLRKQMSIIVTGPVKQKQNFIDLSVASLSAAVSNLVLKILKDTAIDLETQ
KFGVLDVASKRWLIKPTAKSHANGVETHEARKVHVALLEYDEQIVTCDWRRAVAVS
ESVYSDMAKILRLRDRGEPHVSNAKVLVDVGVGCGTKTEILSRVNFDEDLILV
PGQAAMIRRRANSGLIIVATKDNVTVDSPMNFEGKTRCQPKRLFIDEGMLHTG
CVNFLVMSLCDVAYVIGDTQQIPIYNRVSGFFYPAHFSKLEVEDETRITTLRCPAD
VTHYLNRRYEGFVSTSSVKSVSQEMVGAVINPISKPLHGLKILFTTQSDKEALLS
RGYSEVHTVEVQGETYSVSLVRLTPTPISIIAGDSPHVLVALSRHTCSLKYYTVVM
DPLVSIIRDLKLSYLLDMYKVDAGIQ"
4903..5709
/codon_start=1
/product="30K protein"
/protein_id="BAA09877.1"
/db_xref="GI:1619997"
/translation="MAYTQTATTSALLDVTVRGNNSLVNDLAKRLRYDTAVDEFNARDR
RPKVNFSKVISBEBQTLIATRAYPEFQITFYNTQNAVHSLAGLSRSLSELYLMMQIPYG
SLTYDIGNPASHLFGKRAYVHCCMPNLDVRDLMRHEGQKDSIELYSLRDRGKGTVP
NFOKEAPDRYAEPEPDAVCHNTQTCEHQPQQSQKYYAIALHSIVDIPADEFGAALL
RKNVTCYAAPHFSENLLEDSSVNLDEINACFSRDGDKLTFFSASSTLYNCHFSFN
ILKYVCKTYPPASNRREYVYMKFVLVNTVTFCKFSRIDTFLYKGVAAHKSVDSSQFYT
AMEDAWHYKKTLCAMCNSERILLEDDSSVNTWFFPKRDMVIIPLPDI SLETSKRKEV
LVSKDQFVTLNHIPTYQAKALTYANLVSEFVSIRSRVINGVTARSEMDVDKSLQS
LSMTFFLHTKLAVLKDDLLISKFSGLSKTKVCOHWDEISLAFGNAPPSVKERLANKL
IRAAGDALEIKVPDLVITFHDRLVAEYKSSVDMPALDIRKMEETEVMYNALSLSL
RESKDFDVFSQCKSLEVDPMTAAKVIIVAVMSNESGLTILTPERPEANVAQAALDQ
EKASGALVVTSSREVPEEPKQSMGELQLAGFAGHPSSYRNKEIEISLEGFHMA
TADSLRKQMSIIVTGPVKQKQNFIDLSVASLSAAVSNLVLKILKDTAIDLETQ
KFGVLDVASKRWLIKPTAKSHANGVETHEARKVHVALLEYDEQIVTCDWRRAVAVS
ESVYSDMAKILRLRDRGEPHVSNAKVLVDVGVGCGTKTEILSRVNFDEDLILV
PGQAAMIRRRANSGLIIVATKDNVTVDSPMNFEGKTRCQPKRLFIDEGMLHTG
CVNFLVMSLCDVAYVIGDTQQIPIYNRVSGFFYPAHFSKLEVEDETRITTLRCPAD
VTHYLNRRYEGFVSTSSVKSVSQEMVGAVINPISKPLHGLKILFTTQSDKEALLS
RGYSEVHTVEVQGETYSVSLVRLTPTPISIIAGDSPHVLVALSRHTCSLKYYTVVM
DPLVSIIRDLKLSYLLDMYKVDAGIQ"
4903..5709
/codon_start=1
/product="30K protein"
/protein_id="BAA09878.1"
/db_xref="GI:1619998"
/translation="MALVVKGVNINEFIVLTGMEKILPSMTFVKVSVMSKVDKIMV
HENESISEVNLKGVKILDSGVYCLAGLVLTGEMNLPNCRGVSVCLVDKMRERADE
ATLGSYYTAAKKRFQPKVVPVYALITQDAMKNVQVLIINRVMSAGCPLESEFV
SVPCIVYNNIKIGLRKLTNRDGGPELVEEVDVDFMEDVPMISIRLAKRFSRTGKKS
VVPKGNFSRRDSQPNKNYNAKDFGMSFKNNLIDDGSETSVAESDSF"
5712..6191
/codon_start=1
/product="coat protein"
/protein_id="BAA09879.1"
/db_xref="GI:1619999"
/translation="MSYNTNTPSQFVFLSSAWADPIELINLTNALGNQFQTOQARTV
VQRFQSEVWKPSQVTVRFPDSDFKVYRVNAVDPLVLTALIGAPDTNRIIEVENQAN
PSTAETILDRVDDATVAIRSAINNLIIVELTRGTGSYNRSSSESSSGLVMTSSPAT"

```

ORIGIN

Query Match 57.9%; Score 44; DB 14; Length 6395;
Best Local Similarity 73.7%; Fred. No. 0.00031;
Matches 56; Conservative 0; Mismatches 20; Indels 0; Gaps 0;
Qy 1 GTGACAGACGGCTCGCCAAATTGAACTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATCAA 60
Db 5443 GTGAGACGCGGGGTCCCATGGAACCTTACAGAGAGAGTTGTTGATGAGTTTCATGGAAGAT 5502
Qy 61 GTACCAATGGCTGTGA 76
Db 5503 GTACCTATGTCAATTA 5518

Search completed: January 17, 2005, 12:56:39
Job time : 211.545 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2005, 21:45:12 ; Search time 22.7728 Seconds
(without alignments)
17518.955 Million cell updates/sec

Title: US-09-551-494-5_COPY_5430_5505

Perfect score: 76

Sequence: 1 gtgacagacgctgcgaat.....tgaagtaccatgctgtga 76

Scoring table: IDENTITY_NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : N Geneseqn_23Sep04:*

- 1: Geneseqn1980s:*
- 2: Geneseqn1990s:*
- 3: Geneseqn2000s:*
- 4: Geneseqn2001as:*
- 5: Geneseqn2001bs:*
- 6: Geneseqn2002as:*
- 7: Geneseqn2002bs:*
- 8: Geneseqn2003as:*
- 9: Geneseqn2003bs:*
- 10: Geneseqn2003cs:*
- 11: Geneseqn2003ds:*
- 12: Geneseqn2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	6355	3 AAC62372	AAC62372 cDNA sequ
2	76	100.0	6355	9 ACC85005	Acc85005 TMV-U2 ge
3	69.6	91.6	769	11 ADI26338	Adi26338 Novel end
4	69.6	91.6	769	11 ADM68451	Adm68451 Tobacco m
5	69.6	91.6	769	12 ADP26604	Adp26604 Tobamovir
6	68	89.5	772	10 ADI26344	Adi26344 Novel end
7	68	89.5	772	11 ADM68457	Adm68457 Mosaic vi
8	68	89.5	772	12 ADP26610	Adp26610 Heterodup
9	58.6	77.1	411	3 AAC62379	AAC62379 Origin of
10	48.8	64.2	769	10 ADI26343	Adi26343 Novel end
11	48.8	64.2	769	11 ADM68456	Adm68456 Mosaic vi
12	48.8	64.2	769	12 ADP26609	Adp26609 Heterodup
13	45.8	60.3	5997	2 AAQ12188	Aaq12188 Odontoglo
14	45.8	60.3	6597	2 AAQ38106	Aaq38106 ORSV cDNA
15	44	57.9	356	1 AAN30116	Aan30116 TMV-RNA f
16	44	57.9	1019	1 AAN30115	Aan30115 TMV-RNA f
17	42.4	55.8	557	10 ADD17931	Add17931 DNA (Seqi
18	42.4	55.8	557	10 ADK56952	Adk56952 Plant DNA
19	42.4	55.8	632	2 AAV16847	Av16847 Tobacco m
20	42.4	55.8	792	2 ADI26342	Adi26342 Novel end
21	42.4	55.8	792	11 ADM68455	Adm68455 Mosaic vi

C	22	42.4	55.8	792	12	ADP26608	Adp26608 Heterodup
	23	42.4	55.8	804	12	ADN63297	Adn63297 Tobamovir
	24	42.4	55.8	805	6	ABT06582	Abt06582 Wild-type
	25	42.4	55.8	807	6	ABT06572	Abt06572 Wild-type
	26	42.4	55.8	807	8	ABX14577	Abx14577 Tomato mo
	27	42.4	55.8	807	11	ADM68440	Adm68440 Tobacco m
	28	42.4	55.8	807	12	ADP26593	Adp26593 Tobamovir
	29	42.4	55.8	895	2	AAQ62653	Aaq62653 Tobacco-m
	30	42.4	55.8	1187	10	ADD17932	Add17932 DNA (Seqi
	31	42.4	55.8	1187	10	ADK56953	Adk56953 Plant DNA
	32	42.4	55.8	1825	2	AAQ67663	Aaq67663 TMV repli
	33	42.4	55.8	1825	2	AAV54825	Aav54825 Replicon
	34	42.4	55.8	1971	3	AAC62382	Aac62382 cDNA sequ
	35	42.4	55.8	5484	8	ADA15011	Ada15011 Tobacco m
	36	42.4	55.8	6395	2	AAQ51555	Aaq51555 Tobacco m
	37	42.4	55.8	6395	2	AAZ20642	Aaz20642 TMV-based
	38	42.4	55.8	6395	3	AAC62369	Aac62369 cDNA sequ
	39	42.4	55.8	6395	4	AAF82330	Aaf82330 Wild-type
	40	42.4	55.8	6425	2	AAZ20644	Aaz20644 TMV-based
	41	42.4	55.8	6425	4	AAF82332	Aaf82332 Tobacco m
	42	42.4	55.8	6439	2	AAZ20643	Aaz20643 TMV-based
	43	42.4	55.8	6439	4	AAF82331	Aaf82331 Tobacco m
	44	42.4	55.8	6446	2	AAZ20646	Aaz20646 TMV-based
	45	42.4	55.8	6446	4	AAF82334	Aaf82334 Tobacco m

ALIGNMENTS

RESULT 1
AAC62372
ID AAC62372 standard; DNA; 6355 BP.
XX
AC
AAC62372;

19-MAR-2001 (first entry)

cDNA sequence of the genome of tobacco mosaic virus-U2.

Plant phenotype; gene trait; Nicotiana; Oryza sativa; Zea mays; Brassica;
Gossypium; Triticum; Arabidopsis; Petunia; herbicide; transgenic plant;
tobacco necrosis virus; TMV; tobacco mosaic virus; TMV; helper virus; ss.

Tobacco mosaic virus.

WO200063397-A2.

26-OCT-2000.

17-APR-2000; 2000WO-EF003521.

20-APR-1999; 99US-00294022.

(AVET) AVENTIS CROPS SCIENCE NV.

Meulawaeter F, Cornelissen M, Jacobs J, Van Eldik G, Metzlaiff M;

WPI; 2000-687182/67.

Identifying and isolating genes involved in determining the trait or phenotype of plant species, by infecting plants with gene silencing constructs targeted to the gene, and identifying plants with altered traits.

Example 1; Page 53-56; 64pp; English.

The specification describes a method for isolating genes that determine a trait or phenotype of a plant species. The method comprises identifying a set of nucleic acids of genes correlated with the trait, creating a library of gene silencing constructs in a viral RNA vector, targeting the gene silencing constructs to the nucleic acid set, infecting a collection of individual plants with these, identifying plants with altered traits or phenotype, and isolating genes of the invention. The method is useful

for isolating genes involved in the determination of trait or a phenotype of a plant such as Nicotiana, Oryza sativa, Zea mays, Brassica, Gossypium, Triticum, Arabidopsis or Petunia. The method is also useful for modulating the expression of selected nucleic acid sequences and for validating the function of a nucleic acid sequence whose expression is correlated with the presence or absence of a specific trait in plants, but with otherwise unknown function. The method is also useful for developing agronomically useful products such herbicides or transgenic plants. The present sequence represents the cDNA sequence of the genome of tobacco mosaic virus (TMV)-U2. The sequence was used to construct a plasmid vector for the synthesis of an infective hybrid tobacco mosaic virus (TMV)/satellite tobacco necrosis virus (STNV) helper virus RNA. This helper virus is used in the method of the invention

XX SQ Sequence 6355 BP; 1933 A; 1112 C; 1489 G; 1821 T; 0 U; 0 Other;

Query Match 100.0%; Score 76; DB 3; Length 6355;
Best Local Similarity 100.0%; Pred. No. 2.2e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGACAGCGCTCGCCCAATTGAACTCACTGAAAAGGTTGTGAGGAGTTCTGTGGATGAA 60
Db 5430 GTGACAGCGCTCGCCCAATTGAACTCACTGAAAAGGTTGTGAGGAGTTCTGTGGATGAA 5489

Qy 61 GTACCAATGGCTGTGA 76
Db 5490 GTACCAATGGCTGTGA 5505

RESULT 2
ACC85005
ID ACC85005 standard; DNA; 6355 BP.
XX AC ACC85005;
XX AC
XX AC
DT 13-OCT-2003 (first entry)
XX
DE TMV-U2 genome nucleotide sequence.
XX
KW Inhibitory RNA; viral RNA vector; coat protein; TMV; U2; gene; ds.
XX
OS Tobacco mosaic virus.
XX
PN WO2003052108-A2.
XX
PD 26-JUN-2003.
XX
PF 05-DEC-2002; 2002WO-EP013964.
XX
PR 18-DEC-2001; 2001US-0340488P.
XX
PA (FARB) BAYER BIOSCIENCE NV.
XX
PI Metzlaiff MH, Gossele VML, Meulewaeter F, Fache ICA;
XX
DR WPI; 2003-523529/49.

Introducing inhibitory RNA into a plant cell comprises providing a viral RNA vector derived from a satellite RNA virus that encodes a coat protein, and infecting a plant with the viral RNA vector and a corresponding helper virus.

XX Example; Page 79-82; 86pp; English.

XX The invention relates to introducing inhibitory RNA into a plant cell. The method involves providing a viral RNA vector derived from a satellite RNA virus having a sequence that encodes a coat protein, and infecting a plant with the viral RNA vector and a corresponding helper virus. The methods and viral RNA vectors are useful in introducing inhibitory RNA into plant cells. These may be used to determine or validate the function of isolated nucleic acid sequences in plants. The present sequence represents the nucleotide sequence of the genome of tobacco mosaic virus (TMV)-U2

XX SQ Sequence 6355 BP; 1933 A; 1112 C; 1489 G; 1821 T; 0 U; 0 Other;

Query Match 100.0%; Score 76; DB 9; Length 6355;
Best Local Similarity 100.0%; Pred. No. 2.2e-16;
Matches 76; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1 GTGACAGCGCTCGCCCAATTGAACTCACTGAAAAGGTTGTGAGGAGTTCTGTGGATGAA 60
Db 5430 GTGACAGCGCTCGCCCAATTGAACTCACTGAAAAGGTTGTGAGGAGTTCTGTGGATGAA 5489

Qy 61 GTACCAATGGCTGTGA 76
Db 5490 GTACCAATGGCTGTGA 5505

RESULT 3
ADI26338
ID ADI26338 standard; DNA; 769 BP.
XX AC ADI26338;
XX AC
DT 22-APR-2004 (first entry)
XX
DE Novel endonuclease Res I-related clone DNA 3.
XX
KW endonuclease; molecular biology; plant propagation; phenotypic trait; herbicide tolerance; heat tolerance; cold tolerance; drought; salinity; osmotic stress; pest resistance; insect; nematode; arachnid; fungal; bacterial; viral; enzyme production; secondary metabolite; male sterility; female sterility; dwarfness; early maturity; Res I; ds.
XX
OS Tobacco mosaic virus.
XX
PN US2003148315-A1.
XX
PD 07-AUG-2003.
XX
PF 01-AUG-2002; 2002US-00211079.
XX
PR 01-FEB-2002; 2002US-0353722P.
XX
PR 14-MAR-2002; 2002US-00098155.
XX
PA (PADG/) PADGETT H S.
XX
PA (VAEW/) VAETHONGS A A.
XX
PI Padgett HS, Vaethongs AA;
XX
DR WPI; 2003-897548/82.
XX
PT New nucleic acid molecule encoding endonucleases, useful in molecular biology, specifically to generating populations of related nucleic acid molecules, and in plant propagation with useful phenotypic traits.
XX
PS Example 15; Fig 8; 46pp; English.

XX This invention relates to a novel endonuclease (Res I) nucleic acid molecule which comprises a fully defined sequence of 899 bp given in the specification. The methods and compositions of the present invention are useful in molecular biology, and more specifically to generating populations of related nucleic acid molecules. They may also be used in plant propagation with useful phenotypic traits, such as improved tolerance to herbicides, improved tolerance to extremes of heat or cold, drought, salinity or osmotic stress, improved resistance to pests (insects, nematodes or arachnids) or diseases (fungal, bacterial or viral), production of enzymes or secondary metabolites, male or female sterility, dwarfness and early maturity. The present sequence is that of a clone which was derived during the exemplification of the invention.

XX SQ Sequence 769 BP; 247 A; 102 C; 201 G; 219 T; 0 U; 0 Other;

Query Match 91.6%; Score 69.6; DB 10; Length 769;
Best Local Similarity 94.7%; Pred. No. 2.3e-14;

Matches 72; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTGACAGCGCTCGCAATTGAACCTCACTGAAAAGTTGTTGAGGAGTTCTGGATGAA 60
DB 541 GTACACAGCGCTCGCAATTGAACCTCACTGAAAAGTTGTTGAGGAGTTCTAGATGAA 600
QY 61 GTACCAATGGCTGTGA 76
DB 601 GTACCAATGGCTGTGA 616

RESULT 4
ADM68451
ID ADM68451 standard; DNA; 769 BP.
XX
AC ADM68451;
DT 03-JUN-2004 (first entry)
XX
DE Tobacco mosaic virus movement protein gene #4.
XX
KW ds; mismatch endonuclease; endonuclease; gene shuffling technology;
KW single nucleotide polymorphism; cancer susceptibility;
KW sequence variation redistribution; movement protein; gene.
XX
OS Tobacco mosaic virus.
XX
PN US2003157682-A1.
XX
PD 21-AUG-2003.
XX
PF 31-JAN-2003; 2003US-00356708.
XX
PR 01-FEB-2002; 2002US-0353722P.
PR 14-MAR-2002; 2002US-00098155.
PR 01-AUG-2002; 2002US-00211079.
XX
PA (PADG/) PADGETT H S.
PA (VAEW/) VAETHONGS A A.
PA (VOJD/) VOJDANI F S.
PA (SMIT/) SMITH M L.
PA (LIND/) LINDBO J A.
PA (FITZ/) FITZMAURICE W P.

XX Padgett HS, Vaethongs AA, Vojdani FS, Smith ML, Lindbo JA;
PI Fitzmaurice WP;
XX WPI; 2003-766176/72.
XX Making a mismatch endonuclease, useful in gene shuffling and in detection
PT of single nucleotide polymorphisms, comprises transfecting a host with a
PT recombinant viral vector including a polynucleotide encoding a mismatch
PT endonuclease.
XX
PS Example 14; SEQ ID NO 20; 79pp; English.
XX The invention relates to a method of making a mismatch endonuclease
CC enzyme comprising transfecting a host plant, animal, yeast, fungus or
CC bacterium with a recombinant viral vector that encodes a polynucleotide
CC sequence for a mismatch endonuclease, growing the host so that the
CC polynucleotide is expressed, and extracting the mismatch endonuclease
CC enzyme from the host. The method is useful for making mismatch
CC endonuclease enzymes, for obtaining peptides and polynucleotides with
CC desired functional properties and for detecting mutations. The mismatch
CC endonuclease enzymes are useful in gene shuffling technology for
CC developing new genes, in detecting single nucleotide polymorphisms for
CC e.g. detecting evidence of cancer susceptibility, or in redistributing
CC sequence variations between non-identical polynucleotide sequences. The
CC present sequence represents a tobacco mosaic virus movement protein gene.

XX Sequence 769 BP; 247 A; 102 C; 201 G; 219 T; 0 U; 0 Other;
SQ Query Match 91.6%; Score 69.6; DB 11; Length 769;

Best Local Similarity 94.7%; Pred. No. 2.3e-14;
Matches 72; Conservative 0; Mismatches 4; Indels 0; Gaps 0;
QY 1 GTGACAGCGCTCGCAATTGAACCTCACTGAAAAGTTGTTGAGGAGTTCTGGATGAA 60
DB 541 GTACACAGCGCTCGCAATTGAACCTCACTGAAAAGTTGTTGAGGAGTTCTAGATGAA 600
QY 61 GTACCAATGGCTGTGA 76
DB 601 GTACCAATGGCTGTGA 616

RESULT 5
ADP26604
ID ADP26604 standard; DNA; 769 BP.
XX
AC ADP26604;
DT 26-AUG-2004 (first entry)
XX
DE Tobamovirus DNA #5.
XX
KW Sequence variation; heteroduplex; transcription; DNA integration;
KW ribozyme expression; gene; ds; tobamovirus.
XX
OS Tobamovirus.
XX
PN US2004110130-A1.
XX
PD 10-JUN-2004.
XX
PF 25-OCT-2002; 2002US-00280913.
XX
PR 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0268785P.
PR 01-FEB-2002; 2002US-00066390.
PR 08-AUG-2002; 2002US-0402342P.
XX
PA (LARG-) LARGE SCALE BIOLOGY CORP.

XX Padgett HS, Lindbo JA, Fitzmaurice WP;
PI WPI; 2004-440326/41.
XX
XX Redistributing sequence variations between non-identical polynucleotide
PT sequences, useful for generating improved polynucleotide having a desired
PT characteristic, comprises making a heteroduplex and introducing a nick.
XX
PS Example 15; SEQ ID NO 20; 75pp; English.
XX The invention relates to an in vitro method of redistributing sequence
CC variations between non-identical polynucleotide sequences, comprising
CC making a heteroduplex polynucleotide from two non-identical
CC polynucleotides, introducing a nick in the second strand at or near a
CC base pair mismatch site, removing the mismatched base(s) from the
CC mismatch site where the nick occurred and using the first strand as a
CC template to replace the removed base(s) with bases that complement the
CC base(s) in the first strand. The invention also relates to an in vitro
CC method of making a population of sequence variants from a heteroduplex
CC polynucleotide sequence, obtaining a polynucleotide sequence encoding a
CC desired functional property and identifying a reassorted DNA molecule
CC encoding a protein with a desired functional property. The method is
CC useful for generating an improved polynucleotide sequence or a population
CC of improved polynucleotide sequences possessing at least one desired
CC phenotypic characteristic (e.g., promotes transcription of linked
CC polynucleotides), where such polynucleotides are useful for expression
CC from a plant, animal, fungal, yeast, or bacterial expression vector, for
CC integration to form a transgenic plant, animal or microorganism, and for
CC expression of a ribozyme. This sequence represents DNA used in the scope
CC of the invention.

XX Sequence 769 BP; 247 A; 102 C; 201 G; 219 T; 0 U; 0 Other;

Query Match 91.6%; Score 69.6; DB 12; Length 769;
 Best Local Similarity 94.7%; Pred. No. 2.3e-14;
 Matches 72; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

QY 1 GTGACAGACGGCTCGCCAAATGAACTCACTGAAAAAGTGTGTGAGGAGTTCGTGATGAA 60
 |||||
 Db 232 GTAACAGACGGCTCGCCAAATGAACTCACTGAAAAAGTGTGTGAGGAGTTCATAGATGAA 173
 |||||

QY 61 GTACCAATGGCTGTGA 76
 |||||
 Db 172 GTACCAATGGCTGTGA 157
 |||||

RESULT 7
 ADM68457/c
 ID ADM68457 standard; DNA; 772 BP.
 XX
 AC ADM68457;
 XX
 DT 03-JUN-2004 (first entry)
 XX
 DE Mosaic virus movement protein gene GRAMMR clone #11.
 XX
 KW ds; mismatch endonuclease; endonuclease; gene shuffling technology;
 KW single nucleotide polymorphism; cancer susceptibility;
 KW sequence variation redistribution; movement protein; gene.
 XX
 OS Tobacco mosaic virus.
 OS Tomato mosaic virus.
 FN US2003157682-A1.
 XX
 PD 21-AUG-2003.
 XX
 PF 31-JAN-2003; 2003US-00356708.
 XX
 PR 01-FEB-2002; 2002US-0353722P.
 PR 14-MAR-2002; 2002US-00098155.
 PR 01-AUG-2002; 2002US-00211079.
 XX
 PA (PADG/) PADGETT H S.
 PA (VAEW/) VAEWHONGS A A.
 PA (VOJD/) VOJDANI F S.
 PA (SMIT/) SMITH M L.
 PA (LIND/) LINDBO J A.
 PA (FITZ/) FITZMAURICE W P.
 XX
 PI Padgett HS, Vaewhongs AA, Vojdani FS, Smith ML, Lindbo JA;
 PI Fitzmaurice WP;
 XX
 DR WPI; 2003-766176/72.
 XX
 PT Making a mismatch endonuclease, useful in gene shuffling and in detection
 PT of single nucleotide polymorphisms, comprises transfecting a host with a
 PT recombinant viral vector including a polynucleotide encoding a mismatch
 PT endonuclease.
 XX
 PS Example 14; SEQ ID NO 26; 79pp; English.
 XX
 CC The invention relates to a method of making a mismatch endonuclease
 CC enzyme comprising transfecting a host plant, animal, yeast, fungus or
 CC bacterium with a recombinant viral vector that encodes a polynucleotide
 CC sequence for a mismatch endonuclease, growing the host so that the
 CC polynucleotide is expressed, and extracting the mismatch endonuclease
 CC enzyme from the host. The method is useful for making mismatch
 CC endonuclease enzymes, for obtaining peptides and polynucleotides with
 CC desired functional properties and for detecting mutations. The mismatch
 CC endonuclease enzymes are useful in gene shuffling technology for
 CC developing new genes, in detecting single nucleotide polymorphisms for
 CC e.g. detecting evidence of cancer susceptibility, or in redistributing
 CC sequence variations between non-identical polynucleotide sequences. The
 CC present sequence represents a mosaic virus movement protein gene GRAMMR
 CC clone.
 XX
 SQ Sequence 772 BP; 227 A; 196 C; 109 G; 240 T; 0 U; 0 Other;
 Query Match 89.5%; Score 68; DB 11; Length 772;

QY 1 GTGACAGACGGCTCGCCAAATGAACTCACTGAAAAAGTGTGTGAGGAGTTCGTGATGAA 60
 |||||
 Db 541 GTAACAGACGGCTCGCCAAATGAACTCACTGAAAAAGTGTGTGAGGAGTTCATAGATGAA 600
 |||||

QY 61 GTACCAATGGCTGTGA 76
 |||||
 Db 601 GTACCAATGGCTGTGA 616
 |||||

RESULT 6
 ADI26344/c
 ID ADI26344 standard; DNA; 772 BP.
 XX
 AC ADI26344;
 XX
 DT 22-APR-2004 (first entry)
 XX
 DE Novel endonuclease Res I-related clone DNA 9.
 XX
 KW endonuclease; molecular biology; plant propagation; phenotypic trait;
 KW herbicide tolerance; heat tolerance; cold tolerance; drought; salinity;
 KW osmotic stress; pest resistance; insect; nematode; arachnid; fungal;
 KW bacterial; viral; enzyme production; secondary metabolite;
 KW male sterility; female sterility; dwarfness; early maturity; Res I; ds.
 XX
 OS Tobacco mosaic virus.
 OS Tomato mosaic virus.
 XX
 FN US2003148315-A1.
 XX
 PD 07-AUG-2003.
 XX
 PF 01-AUG-2002; 2002US-00211079.
 XX
 PR 01-FEB-2002; 2002US-0353722P.
 PR 14-MAR-2002; 2002US-00098155.
 XX
 PA (PADG/) PADGETT H S.
 PA (VAEW/) VAEWHONGS A A.
 XX
 PI Padgett HS, Vaewhongs AA;
 XX
 DR WPI; 2003-897548/82.
 XX
 PT New nucleic acid molecule encoding endonucleases, useful in molecular
 PT biology, specifically to generating populations of related nucleic acid
 PT molecules, and in plant propagation with useful phenotypic traits.
 XX
 PS Example 15; Fig 14; 46pp; English.
 XX
 CC This invention relates to a novel endonuclease (Res I) nucleic acid
 CC molecule which comprises a fully defined sequence of 899 bp given in the
 CC specification. The methods and compositions of the present invention are
 CC useful in molecular biology, and more specifically to generating
 CC populations of related nucleic acid molecules. They may also be used in
 CC plant propagation with useful phenotypic traits, such as improved
 CC tolerance to herbicides, improved tolerance to extremes of heat or cold,
 CC drought, salinity or osmotic stress, improved resistance to pests
 CC (insects, nematodes or arachnids) or diseases (fungal, bacterial or
 CC viral), production of enzymes or secondary metabolites, male or female
 CC sterility, dwarfness and early maturity. The present sequence is that of
 CC a clone which was derived during the exemplification of the invention.
 XX
 SQ Sequence 772 BP; 227 A; 196 C; 109 G; 240 T; 0 U; 0 Other;
 Query Match 89.5%; Score 68; DB 10; Length 772;
 Best Local Similarity 93.4%; Pred. No. 8.2e-14;
 Matches 71; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Best Local Similarity 93.4%; Pred. No. 8.2e-14;
Matches 71; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 GTGACAGACGGCTCGCCAAATTGAACCTCACTGAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
DB 232 GTAACAGACGGCTCGCCAAATTGAACCTCACTGAAAGGTTGTTGATGAGTTTCATAGATGAA 173
QY 61 GTACCAATGGCTGTGA 76
DB 172 GTACCAATGGCTGTGA 157

RESULT 8
ADP26610/C
ID ADP26610 standard; DNA; 772 BP.
XX AC ADP26610;
XX AC
DT 26-AUG-2004 (first entry)
XX DE Heteroduplex DNA #15.
XX Sequence variation; heteroduplex; transcription; DNA integration;
KW ribozyme expression; gene; ds.
XX OS Synthetic.
XX US2004110130-A1.
XX 10-JUN-2004.
XX 25-OCT-2002; 2002US-00280913.
XX 02-FEB-2001; 2001US-0266386P.
PR 14-FEB-2001; 2001US-0268785P.
PR 01-FEB-2002; 2002US-00066390.
PR 08-AUG-2002; 2002US-0402342P.
XX (LARG-) LARGE SCALE BIOLOGY CORP.

XX Padgett HS, Lindbo JA, Fitzmaurice WP;
XX WPI; 2004-440326/41.
XX
XX Redistributing sequence variations between non-identical polynucleotide
PT sequences, useful for generating improved polynucleotide having a desired
PT characteristic, comprises making a heteroduplex and introducing a nick.
XX
XX Example 15; SEQ ID NO 26; 75pp; English.

XX The invention relates to an in vitro method of redistributing sequence
CC variations between non-identical polynucleotide sequences, comprising
CC making a heteroduplex polynucleotide from two non-identical
CC polynucleotides, introducing a nick in the second strand at or near a
CC base pair mismatch site, removing the mismatched base(s) from the
CC mismatch site where the nick occurred and using the first strand as a
CC template to replace the removed base(s) with bases that complement the
CC base(s) in the first strand. The invention also relates to an in vitro
CC method of making a population of sequence variants from a heteroduplex
CC polynucleotide sequence, obtaining a polynucleotide sequence encoding a
CC desired functional property and identifying a reassorted DNA molecule
CC encoding a protein with a desired functional property. The method is
CC useful for generating an improved polynucleotide sequence or a population
CC of improved polynucleotide sequences possessing at least one desired
CC phenotypic characteristic (e.g., promotes transcription of linked
CC polynucleotides), where such polynucleotides are useful for expression
CC from a plant, animal, fungal, yeast, or bacterial expression vector, for
CC integration to form a transgenic plant, animal or microorganism, and for
CC expression of a ribozyme. This sequence represents DNA used in the scope
CC of the invention.

XX Sequence 772 BP; 227 A; 196 C; 109 G; 240 T; 0 U; 0 Other;

Query Match 89.5%; Score 68; DB 12; Length 772;
Best Local Similarity 93.4%; Pred. No. 8.2e-14;
Matches 71; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
QY 1 GTGACAGACGGCTCGCCAAATTGAACCTCACTGAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
DB 232 GTAACAGACGGCTCGCCAAATTGAACCTCACTGAAAGGTTGTTGATGAGTTTCATAGATGAA 173
QY 61 GTACCAATGGCTGTGA 76
DB 172 GTACCAATGGCTGTGA 157

RESULT 9
AAC62379
ID AAC62379 standard; DNA; 411 BP.
XX AC AAC62379;
XX AC
DT 19-MAR-2001 (first entry)
XX DE Origin of assembly (OAS) of a tobacco mosaic virus (TMV)-U2.

XX Plant phenotype; gene trait; Nicotiana; Oryza sativa; Zea mays; Brassica;
KW Gossypium; Triticum; Arabidopsis; Petunia; herbicide; transgenic plant;
KW tobacco necrosis virus; TMV; tobacco mosaic virus; TMV; helper virus;
KW origin of assembly; ss.

XX Tobacco mosaic virus.

XX WO200063397-A2.

XX 26-OCT-2000.

XX 17-APR-2000; 2000WO-BP003521.

XX 20-APR-1999; 99US-00294022.

XX (AVET) AVENTIS CROPS SCIENCE NV.

XX Meulewaeter F, Cornelissen M, Jacobs J, Van Eldik G, Metzlaiff M;

XX WPI; 2000-687182/67.

XX Identifying and isolating genes involved in determining the trait or
PT phenotype of plant species, by infecting plants with gene silencing
PT constructs targeted to the gene, and identifying plants with altered
PT traits.

XX Example 1; Page 63; 64pp; English.

XX The specification describes a method for isolating genes that determine a
CC trait or phenotype of a plant species. The method comprises identifying a
CC set of nucleic acids of genes correlated with the trait, creating a
CC library of gene silencing constructs in a viral RNA vector, targeting the
CC gene silencing constructs to the nucleic acid set, infecting a collection
CC of individual plants with these, identifying plants with altered traits
CC or phenotype, and isolating genes of the invention. The method is useful
CC for isolating genes involved in the determination of trait or a phenotype
CC of a plant such as Nicotiana, Oryza sativa, Zea mays, Brassica, Gossypium,
CC Triticum, Arabidopsis or Petunia. The method is also useful for
CC modulating the expression of selected nucleic acid sequences and for
CC validating the function of a nucleic acid sequence whose expression is
CC correlated with the presence or absence of a specific trait in plants,
CC but with otherwise unknown function. The method is also useful for
CC developing agronomically useful products such herbicides or transgenic
CC plants. The present sequence is an origin of assembly (OAS) of a tobacco
CC mosaic virus (TMV)-U2. The sequence is used to construct infective hybrid
CC tobacco mosaic virus (TMV)/tobacco necrosis virus (TMV) vectors, for use
CC in the method of the invention

XX Sequence 411 BP; 140 A; 70 C; 93 G; 108 T; 0 U; 0 Other;


```

;
; ZIP: 94111
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Floppy disk
; COMPUTER: IBM PC compatible
; OPERATING SYSTEM: PC-DOS/MS-DOS
; SOFTWARE: Patent in Release #1.0, Version #1.25
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/336,724
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P.
; REGISTRATION NUMBER: 25,227
; REFERENCE/DOCKET NUMBER: US 07/997,733
; FILING DATE: 30-DEC-1992
; TELEPHONE: 415-433-4150
; TELEFAX: 415-433-8716
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1825
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: RNA (episomal), peptide
; DESCRIPTION: Peptide encodes for TMV 30kDa
; DESCRIPTION: movement protein (268 residues) and CAT (204
; HYPOTHEICAL: NO
; ANTI-SENSE: NO
; ORIGINAL SOURCE:
; ORGANISM: Tobacco Mosaic Virus
; IMMEDIATE SOURCE:
; CLONE:
; FEATURE:
;
; US-08-336-724-1
;
; Query Match 55.8%; Score 42.4; DB 2; Length 1825;
; Best Local Similarity 56.6%; Pred. No. 2.7e-06;
; Matches 43; Conservative 12; Mismatches 21; Indels 0; Gaps 0;
;
; QY 1 GTGACAGCGCTCGCCAAATTAACCTCACTGAAAGGTTGTGAGGAGTTGTCGTGATGAA 60
; Db 610 GUGAGAGCGAGGCGCCCGAUGAAGCUACAGAGAGTGTGUGAUGAGUUGAAGAU 669
;
; QY 61 GTACCAATGGCTGTGA 76
; Db 670 GUCCUAUGUGCAUCA 685
;
; RESULT 6
; US-08-687-559-2
; Sequence 2, Application US/08687559
; Patent No. 5955647
; GENERAL INFORMATION:
; APPLICANT: Fitch, John H.
; APPLICANT: Beachy, Roger N.
; TITLE OF INVENTION: METHOD FOR USING TOBACCO MOSAIC VIRUS TO
; TITLE OF INVENTION: OVERPRODUCE PEPTIDES AND PROTEINS
; NUMBER OF SEQUENCES: 30
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Fish & Richardson, P.C.
; STREET: 4225 Executive Square, Suite 1400
; CITY: La Jolla
; STATE: CA
; COUNTRY: US
; ZIP: 92037
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: Windows95
;
; US-08-687-559-2
; Sequence 1, Application US/09259741
; Patent No. 6033895
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; APPLICANT: HOLTZ, R. BARRY
; APPLICANT: McCULLOCH, MICHAEL
; APPLICANT: TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT
; TITLE OF INVENTION: SOURCES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/259,741
; FILING DATE: February 25, 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751
; FILING DATE: March 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
;
; US-08-687-559-2
; Query Match 55.8%; Score 42.4; DB 2; Length 6395;
; Best Local Similarity 72.4%; Pred. No. 4.3e-06;
; Matches 55; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
;
; QY 1 GTGACAGCGCTCGCCAAATTAACCTCACTGAAAGGTTGTGAGGAGTTGTCGTGATGAA 60
; Db 5443 GTGAGAGCGAGGCGCCCGAUGAAGCUACAGAGAGTGTGUGAUGAGUUGAAGAU 5502
;
; QY 61 GTACCAATGGCTGTGA 76
; Db 5503 GTCCCTATGTCGATCA 5518
;
; RESULT 7
; US-09-259-741-1
; Sequence 1, Application US/09259741
; Patent No. 6033895
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; APPLICANT: HOLTZ, R. BARRY
; APPLICANT: McCULLOCH, MICHAEL
; APPLICANT: TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT
; TITLE OF INVENTION: SOURCES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/259,741
; FILING DATE: February 25, 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751
; FILING DATE: March 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
;
; US-08-687-559-2
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/08/687,559
; FILING DATE: No. 5955647ember 18, 1996
; CLASSIFICATION: 424
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: PCT/US95/01467
; FILING DATE: 03-FEB-1995
; ATTORNEY/AGENT INFORMATION:
; NAME: Learn, June M.
; REGISTRATION NUMBER: 31,238
; REFERENCE/DOCKET NUMBER: 07302/011001
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 619-678-5070
; TELEFAX: 619-678-5099
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: linear
; MOLECULE TYPE: Genomic DNA
; IMMEDIATE SOURCE:
; CLONE: TMV
;
; US-08-687-559-2
```

TGAA 60
||


```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/259,741
; FILING DATE: February 25, 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751
; FILING DATE: March 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
; REFERENCE/DOCKET NUMBER: 00801.0140.US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8100
; TELEFAX: 650-463-8400
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6425 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: Genomic RNA
; US-09-259-741-3

Query Match 55.8%; Score 42.4; DB 3; Length 6425;
Best Local Similarity 56.6%; Pred. No. 4.3e-06;
Matches 43; Conservative 12; Mismatches 21; Indels 0; Gaps 0;

Qy 1 GTGACAGACGGCTCGCCAAATTGAACCTCACTGAAAAGTTGTGTGAGGAGTTCGTGGATGAA 60
Db 5443 GUGAGAGACGGAGGGCCCAUGGAACUACAGAGAGUCGUGAUGAUGAUGAUGAUGAUGAUGAUG 5502
Qy 61 GTACCAATGGCTGTGA 76
Db 5503 GUCCCUAUGUGCAUCA 5518

RESULT 13
US-09-037-751-3
; Sequence 3, Application US/09037751
; Patent No. 6037456
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; APPLICANT: HOLTZ, R. BARRY
; APPLICANT: MCCULLOCH, MICHAEL
; APPLICANT: TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/466,422
; FILING DATE: 17-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277

```

```

; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/037,751
; FILING DATE: 10-MAR-1998
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER:
; FILING DATE:
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
; REFERENCE/DOCKET NUMBER: 00801.0140.999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8109
; TELEFAX: 650-463-8400
; TELEX:
; INFORMATION FOR SEQ ID NO: 3:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6425 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: Genomic RNA
; US-09-037-751-3

Query Match 55.8%; Score 42.4; DB 3; Length 6425;
Best Local Similarity 56.6%; Pred. No. 4.3e-06;
Matches 43; Conservative 12; Mismatches 21; Indels 0; Gaps 0;

Qy 1 GTGACAGACGGCTCGCCAAATTGAACCTCACTGAAAAGTTGTGTGAGGAGTTCGTGGATGAA 60
Db 5443 GUGAGAGACGGAGGGCCCAUGGAACUACAGAGAGUCGUGAUGAUGAUGAUGAUGAUGAUGAUG 5502
Qy 61 GTACCAATGGCTGTGA 76
Db 5503 GUCCCUAUGUGCAUCA 5518

RESULT 14
US-09-466-422-3
; Sequence 3, Application US/09466422
; Patent No. 6303779
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; APPLICANT: HOLTZ, R. BARRY
; APPLICANT: MCCULLOCH, MICHAEL
; APPLICANT: TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/466,422
; FILING DATE: 17-Dec-1999
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751
; FILING DATE: <Unknown>
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277

```

REFERENCE/DOCKET NUMBER: 00801.0140.999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-463-8109
TELEFAX: 650-463-8400
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6425 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: Genomic RNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-466-422-3
Query Match 55.8%; Score 42.4; DB 3; Length 6425;
Best Local Similarity 56.6%; Pred. No. 4.3e-06;
Matches 43; Conservative 12; Mismatches 21; Indels 0; Gaps 0;
Qy 1 GTGACAGACGGCTCGCCAAATTGAACTCACTCAAAAGTTTGTGAGGAGTTCTGTCGATGAA 60
Db 5443 GUGAGAGACGGAGGGCCCAUGGAACUACAGAGAGGUCGUUGAUGAGUUCUUGGAAGAU 5502
Qy 61 GTACCAATGGCTGTGA 76
Db 5503 GUCCCUAUGUGCAUCA 5518
Search completed: January 17, 2005, 16:36:12
Job time : 5.08894 secs
TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
FROM PLANT SOURCES
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
ADDRESSEE: Howrey & Simon
STREET: 1299 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/962,527
FILING DATE: 24-Sep-2001
CLASSIFICATION: <Unknown>
PRIORITY APPLICATION DATA:
APPLICATION NUMBER: 09/037,751
FILING DATE: 10-march-1998
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P
REGISTRATION NUMBER: 25,277
REFERENCE/DOCKET NUMBER: 00801.0140.999
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-463-8109
TELEFAX: 650-463-8400
INFORMATION FOR SEQ ID NO: 3:
SEQUENCE CHARACTERISTICS:
LENGTH: 6425 base pairs
TYPE: nucleic acid
STRANDEDNESS: single

TOPOLOGY: unknown
MOLECULE TYPE: Genomic RNA
SEQUENCE DESCRIPTION: SEQ ID NO: 3:
US-09-962-527-3
Query Match 55.8%; Score 42.4; DB 4; Length 6425;
Best Local Similarity 56.6%; Pred. No. 4.3e-06;
Matches 43; Conservative 12; Mismatches 21; Indels 0; Gaps 0;
Qy 1 GTGACAGACGGCTCGCCAAATTGAACTCACTCAAAAGTTTGTGAGGAGTTCTGTCGATGAA 60
Db 5443 GUGAGAGACGGAGGGCCCAUGGAACUACAGAGAGGUCGUUGAUGAGUUCUUGGAAGAU 5502
Qy 61 GTACCAATGGCTGTGA 76
Db 5503 GUCCCUAUGUGCAUCA 5518
Search completed: January 17, 2005, 16:36:12
Job time : 5.08894 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2005, 12:56:43 ; Search time 26.3891 Seconds
(without alignments)
16548.056 Million cell updates/sec

Title: US-09-551-494-5_COPY_5430_5505
Perfect score: 76
Sequence: 1 gtgacagagcgtcgcaat.....tgaagtaccaaagctgtga 76

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 4300275 seqs, 2872944193 residues
Total number of hits satisfying chosen parameters: 8600550

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published Applications NA:
1: /cgn2_6/ptodata/1/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/1/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
4: /cgn2_6/ptodata/1/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/1/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/1/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/1/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/1/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/1/pubpna/US09A_PUBCOMB.seq:*
10: /cgn2_6/ptodata/1/pubpna/US09B_PUBCOMB.seq:*
11: /cgn2_6/ptodata/1/pubpna/US09C_PUBCOMB.seq:*
12: /cgn2_6/ptodata/1/pubpna/US09_NEW_PUB.seq:*
13: /cgn2_6/ptodata/1/pubpna/US10A_PUBCOMB.seq:*
14: /cgn2_6/ptodata/1/pubpna/US10B_PUBCOMB.seq:*
15: /cgn2_6/ptodata/1/pubpna/US10C_PUBCOMB.seq:*
16: /cgn2_6/ptodata/1/pubpna/US10D_PUBCOMB.seq:*
17: /cgn2_6/ptodata/1/pubpna/US10E_PUBCOMB.seq:*
18: /cgn2_6/ptodata/1/pubpna/US10_NEW_PUB.seq:*
19: /cgn2_6/ptodata/1/pubpna/US11_NEW_PUB.seq:*
20: /cgn2_6/ptodata/1/pubpna/US60_NEW_PUB.seq:*
21: /cgn2_6/ptodata/1/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	76	100.0	6355	15	US-10-321-434-7
2	69.6	91.6	769	15	Sequence 7, Appl
3	69.6	91.6	769	15	Sequence 21, Appl
4	69.6	91.6	769	17	Sequence 20, Appl
5	69.6	91.6	769	17	Sequence 20, Appl
6	69.6	91.6	769	17	Sequence 20, Appl
7	68	89.5	772	15	Sequence 27, Appl
8	68	89.5	772	15	Sequence 26, Appl
9	68	89.5	772	15	Sequence 26, Appl
10	68	89.5	772	17	Sequence 26, Appl
11	68	89.5	772	17	Sequence 26, Appl
12	48.8	64.2	769	15	Sequence 26, Appl

c	13	48.8	64.2	769	15	US-10-356-708-25	Sequence 25, Appl
c	14	48.8	64.2	769	17	US-10-280-913A-25	Sequence 25, Appl
c	15	48.8	64.2	769	17	US-10-684-134-25	Sequence 25, Appl
c	16	48.8	64.2	769	17	US-10-637-758-25	Sequence 25, Appl
c	17	42.4	55.8	792	15	US-10-211-079-25	Sequence 25, Appl
c	18	42.4	55.8	792	15	US-10-356-708-24	Sequence 24, Appl
c	19	42.4	55.8	792	17	US-10-280-913A-24	Sequence 24, Appl
c	20	42.4	55.8	792	17	US-10-684-134-24	Sequence 24, Appl
c	21	42.4	55.8	792	17	US-10-637-758-24	Sequence 24, Appl
c	22	42.4	55.8	805	10	US-09-775-049-27	Sequence 27, Appl
c	23	42.4	55.8	807	10	US-09-775-049-17	Sequence 17, Appl
c	24	42.4	55.8	807	13	US-10-066-390-10	Sequence 10, Appl
c	25	42.4	55.8	807	13	US-10-206-030-10	Sequence 10, Appl
c	26	42.4	55.8	807	15	US-10-211-079-10	Sequence 9, Appl
c	27	42.4	55.8	807	15	US-10-356-708-9	Sequence 10, Appl
c	28	42.4	55.8	807	15	US-10-205-772-10	Sequence 2, Appl
c	29	42.4	55.8	807	17	US-10-280-913A-9	Sequence 5, Appl
c	30	42.4	55.8	807	17	US-10-684-134-9	Sequence 5, Appl
c	31	42.4	55.8	807	17	US-10-637-758-9	Sequence 9, Appl
c	32	42.4	55.8	1824	9	US-09-930-329-1	Sequence 1, Appl
c	33	42.4	55.8	1824	9	US-09-930-342-1	Sequence 1, Appl
c	34	42.4	55.8	6395	10	US-09-962-527-1	Sequence 2, Appl
c	35	42.4	55.8	6395	15	US-10-338-092-2	Sequence 1, Appl
c	36	42.4	55.8	6395	17	US-10-828-029-1	Sequence 3, Appl
c	37	42.4	55.8	6425	10	US-09-962-527-3	Sequence 3, Appl
c	38	42.4	55.8	6425	17	US-10-828-029-3	Sequence 2, Appl
c	39	42.4	55.8	6439	10	US-09-962-527-2	Sequence 2, Appl
c	40	42.4	55.8	6439	17	US-10-828-029-2	Sequence 2, Appl
c	41	42.4	55.8	6446	10	US-09-962-527-5	Sequence 5, Appl
c	42	42.4	55.8	6446	17	US-10-828-029-5	Sequence 5, Appl
c	43	42.4	55.8	6475	10	US-09-962-527-4	Sequence 4, Appl
c	44	42.4	55.8	6475	17	US-10-828-029-4	Sequence 4, Appl
c	45	42.4	55.8	7685	9	US-09-949-317-22	Sequence 22, Appl

ALIGNMENTS

RESULT 1
US-10-321-434-7
; Sequence 7, Application US/10321434
; Publication No. US20030135882A1
; GENERAL INFORMATION:
; APPLICANT: Metzlaff, Michael
; APPLICANT: Meulewater, Frank
; APPLICANT: Gossel, Veronique
; APPLICANT: Fach, Ina
; TITLE OF INVENTION: Improved methods and means for delivering inhibitory RNA to plant
; FILE REFERENCE: FKOMOD
; CURRENT APPLICATION NUMBER: US/10/321,434
; CURRENT FILING DATE: 2002-12-18
; NUMBER OF SEQ ID NOS: 18
; SOFTWARE: Patent in version 3.0
; SEQ ID NO 7
; LENGTH: 6355
; TYPE: DNA
; ORGANISM: Artificial
; FEATURE:
; OTHER INFORMATION: cDNA sequence of the genome of TMV-U2
US-10-321-434-7

Qy	1	GTGACAGACGCTCGCCAAATTAACCTCACTGAAAGGTTCTTGAGGATTCGTGATGAA	60
Db	5430	GTGACAGACGCTCGCCAAATTAACCTCACTGAAAGGTTCTTGAGGATTCGTGATGAA	5489
Qy	61	GTACCAATGGCTGTGA	76
Db	5490	GTACCAATGGCTGTGA	5505

```
RESULT 2
US-10-211-079-21
; Sequence 21, Application US/102111079
; Publication No. US20030148315A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Vawhongs, Andrew A.
; APPLICANT: Vojdani, Fakhrieh S.
; APPLICANT: Smith, Mark L.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I
; TITLE OF INVENTION: Endonuclease and Methods of Use Thereof
; FILE REFERENCE: P-LG 5381
; CURRENT APPLICATION NUMBER: US/10/211,079
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 10/098,155
; PRIOR FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 21
; LENGTH: 769
; TYPE: DNA
; ORGANISM: tobamovirus
US-10-211-079-21

Query Match          91.6%; Score 69.6; DB 15; Length 769;
Best Local Similarity 94.7%; Pred. No. 1.4e-14;
Matches 72; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTGACAGCGCTCGCCAAATGAACCTCACTGAAAAGTTGTTGAGGAGTTCTGGATGAA 60
Db 541 GTACACAGCGCTCGCCAAATGAACCTCACTGAAAAGTTGTTGAGGAGTTCTAGATGAA 600

Qy 61 GTACCAATGGCTGTGA 76
Db 601 GTACCAATGGCTGTGA 616

RESULT 3
US-10-356-708-20
; Sequence 20, Application US/10356708
; Publication No. US20030157682A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF
; FILE REFERENCE: P-LG 10100
; CURRENT APPLICATION NUMBER: US/10/356,708
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: 60/353,722
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 10/098,155
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 10/211,079
; PRIOR FILING DATE: 2002-08-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 20
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus-U2
US-10-356-708-20

Query Match          91.6%; Score 69.6; DB 15; Length 769;
Best Local Similarity 94.7%; Pred. No. 1.4e-14;
Matches 72; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTGACAGCGCTCGCCAAATGAACCTCACTGAAAAGTTGTTGAGGAGTTCTGGATGAA 60
Db 541 GTACACAGCGCTCGCCAAATGAACCTCACTGAAAAGTTGTTGAGGAGTTCTAGATGAA 600

Qy 61 GTACCAATGGCTGTGA 76
Db 601 GTACCAATGGCTGTGA 616
```

```
Db 601 GTACCAATGGCTGTGA 616

RESULT 4
US-10-280-913A-20
; Sequence 20, Application US/10280913A
; Publication No. US20040110130A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLGS392-CIP
; CURRENT APPLICATION NUMBER: US/10/280,913A
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 20
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus-U2
US-10-280-913A-20

Query Match          91.6%; Score 69.6; DB 17; Length 769;
Best Local Similarity 94.7%; Pred. No. 1.4e-14;
Matches 72; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTGACAGCGCTCGCCAAATGAACCTCACTGAAAAGTTGTTGAGGAGTTCTGGATGAA 60
Db 541 GTACACAGCGCTCGCCAAATGAACCTCACTGAAAAGTTGTTGAGGAGTTCTAGATGAA 600

Qy 61 GTACCAATGGCTGTGA 76
Db 601 GTACCAATGGCTGTGA 616

RESULT 5
US-10-684-134-20
; Sequence 20, Application US/10684134
; Publication No. US20040142433A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLGS392-CIP
; CURRENT APPLICATION NUMBER: US/10/684,134
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 20
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus-U2
US-10-684-134-20

Query Match          91.6%; Score 69.6; DB 17; Length 769;
Best Local Similarity 94.7%; Pred. No. 1.4e-14;
Matches 72; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTGACAGCGCTCGCCAAATGAACCTCACTGAAAAGTTGTTGAGGAGTTCTGGATGAA 60
```

Db 541 GTACACAGCGCTCGCCCAATTGAACTCACTGAAAAAGTTGTTGAGGAGTTTCATAGTAA 600
Qy 61 GTACCAATGGCTGTGA 76
Db 601 GTACCAATGGCTGTGA 616

RESULT 6
US-10-637-758-20
; Sequence 20, Application US/10637758
; Publication No. US20040180352A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS

; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/637,758
; PRIOR FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US/10/280,913
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 20
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Tobacco mosaic virus-U2
US-10-637-758-20

Query Match 91.6%; Score 69.6; DB 17; Length 769;
Best Local Similarity 94.7%; Pred. No. 1.4e-14;
Matches 72; Conservative 0; Mismatches 4; Indels 0; Gaps 0;

Qy 1 GTGACAGACGCGCTCGCCCAATTGAACTCACTGAAAAAGTTGTTGAGGAGTTTCGTGATGAA 60
Db 541 GTACACAGCGCTCGCCCAATTGAACTCACTGAAAAAGTTGTTGAGGAGTTTCATAGTAA 600
Qy 61 GTACCAATGGCTGTGA 76
Db 601 GTACCAATGGCTGTGA 616

RESULT 7
US-10-211-079-27/c
; Sequence 27, Application US/10211079
; Publication No. US20030148315A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Vaewhongs, Andrew A.
; APPLICANT: Vojdani, Fakhrish S.
; APPLICANT: Smith, Mark L.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I
; FILE REFERENCE: P-LG 5381
; CURRENT APPLICATION NUMBER: US/10/211,079
; PRIOR FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 10/098,155
; PRIOR FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 27
; LENGTH: 772
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct

US-10-211-079-27

Query Match 89.5%; Score 68; DB 15; Length 772;
Best Local Similarity 93.4%; Pred. No. 5.2e-14;
Matches 71; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GTGACAGACGCGCTCGCCCAATTGAACTCACTGAAAAAGTTGTTGAGGAGTTTCGTGATGAA 60
Db 232 GTACACAGCGCTCGCCCAATTGAACTCACTGAAAAAGTTGTTGATGAGTTTCATAGTAA 173
Qy 61 GTACCAATGGCTGTGA 76
Db 172 GTACCAATGGCTGTGA 157

RESULT 8
US-10-356-708-26/c
; Sequence 26, Application US/10356708
; Publication No. US20030157682A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF
; FILE REFERENCE: P-LG 10100
; CURRENT APPLICATION NUMBER: US/10/356,708
; CURRENT FILING DATE: 2003-01-31
; PRIOR APPLICATION NUMBER: 60/353,722
; PRIOR FILING DATE: 2000-02-01
; PRIOR APPLICATION NUMBER: 10/098,155
; PRIOR FILING DATE: 2002-03-14
; PRIOR APPLICATION NUMBER: 10/211,079
; PRIOR FILING DATE: 2002-08-01
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 26
; LENGTH: 772
; TYPE: DNA
; ORGANISM: Artificial Sequence derived from TMV & TMV-U2
; FEATURE:
; OTHER INFORMATION: This sequence was derived by shuffling in accordance with the met
; OTHER INFORMATION: hology of the present invention.
US-10-356-708-26

Query Match 89.5%; Score 68; DB 15; Length 772;
Best Local Similarity 93.4%; Pred. No. 5.2e-14;
Matches 71; Conservative 0; Mismatches 5; Indels 0; Gaps 0;

Qy 1 GTGACAGACGCGCTCGCCCAATTGAACTCACTGAAAAAGTTGTTGAGGAGTTTCGTGATGAA 60
Db 232 GTACACAGCGCTCGCCCAATTGAACTCACTGAAAAAGTTGTTGATGAGTTTCATAGTAA 173
Qy 61 GTACCAATGGCTGTGA 76
Db 172 GTACCAATGGCTGTGA 157

RESULT 9
US-10-280-913A-26/c
; Sequence 26, Application US/10280913A
; Publication No. US20040110130A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/280,913A
; CURRENT FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02

NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 26
; LENGTH: 772
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This construct was derived by GRAMMR shuffling in accordance with
; OTHER INFORMATION: the methodology of the present invention.
US-10-280-913A-26

Query Match 89.5%; Score 68; DB 17; Length 772;
Best Local Similarity 93.4%; Pred. No. 5.2e-14;
Matches 71; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 GTGACAGACGGCTCGCCCAATTGAACCTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
Db 232 GTACACAGCGCTCGCCCAATTGAACCTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 173
Qy 61 GTACCAATGGCTGTGA 76
Db 172 GTACCAATGGCTGTGA 157

RESULT 10
US-10-684-134-26/c
; Sequence 26, Application US/10684134
; Publication No. US20040142433A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS
; FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/684.134
; CURRENT FILING DATE: 2003-10-10
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 26
; LENGTH: 772
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This construct was derived by GRAMMR shuffling in accordance with
; OTHER INFORMATION: the methodology of the present invention.
US-10-684-134-26

Query Match 89.5%; Score 68; DB 17; Length 772;
Best Local Similarity 93.4%; Pred. No. 5.2e-14;
Matches 71; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 GTGACAGACGGCTCGCCCAATTGAACCTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
Db 232 GTACACAGCGCTCGCCCAATTGAACCTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 173
Qy 61 GTACCAATGGCTGTGA 76
Db 172 GTACCAATGGCTGTGA 157

RESULT 11
US-10-637-758-26/c
; Sequence 26, Application US/10637758
; Publication No. US20040180352A1
; GENERAL INFORMATION:
; APPLICANT: Large Scale Biology Corporation
; TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS

FILE REFERENCE: LSBC-PLG5392-CIP
; CURRENT APPLICATION NUMBER: US/10/637,758
; CURRENT FILING DATE: 2003-08-08
; PRIOR APPLICATION NUMBER: US/10/280,913
; PRIOR FILING DATE: 2002-10-25
; PRIOR APPLICATION NUMBER: 60/402,342
; PRIOR FILING DATE: 2002-08-08
; PRIOR APPLICATION NUMBER: 10/066,390
; PRIOR FILING DATE: 2002-02-01
; PRIOR APPLICATION NUMBER: 60/268,785
; PRIOR FILING DATE: 2001-02-14
; PRIOR APPLICATION NUMBER: 60/266,386
; PRIOR FILING DATE: 2002-02-02
; NUMBER OF SEQ ID NOS: 35
; SOFTWARE: Patent in version 3.1
; SEQ ID NO 26
; LENGTH: 772
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: This construct was derived by GRAMMR shuffling in accordance with
; OTHER INFORMATION: the methodology of the present invention.
US-10-637-758-26

Query Match 89.5%; Score 68; DB 17; Length 772;
Best Local Similarity 93.4%; Pred. No. 5.2e-14;
Matches 71; Conservative 0; Mismatches 5; Indels 0; Gaps 0;
Qy 1 GTGACAGACGGCTCGCCCAATTGAACCTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
Db 232 GTACACAGCGCTCGCCCAATTGAACCTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 173
Qy 61 GTACCAATGGCTGTGA 76
Db 172 GTACCAATGGCTGTGA 157

RESULT 12
US-10-211-079-26/c
; Sequence 26, Application US/10211079
; Publication No. US20030148315A1
; GENERAL INFORMATION:
; APPLICANT: Padgett, Hal S.
; APPLICANT: Vaewhongs, Andrew A.
; APPLICANT: Vojdani, Fakhrhien S.
; APPLICANT: Smith, Mark L.
; TITLE OF INVENTION: Nucleic Acid Molecules Encoding CEL I
; TITLE OF INVENTION: Endonuclease and Methods of Use Thereof
; FILE REFERENCE: P-LG 5381
; CURRENT APPLICATION NUMBER: US/10/211,079
; CURRENT FILING DATE: 2002-08-01
; PRIOR APPLICATION NUMBER: US 10/098,155
; PRIOR FILING DATE: 2002-03-14
; NUMBER OF SEQ ID NOS: 29
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 26
; LENGTH: 769
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: synthetic construct
US-10-211-079-26

Query Match 64.2%; Score 48.8; DB 15; Length 769;
Best Local Similarity 77.6%; Pred. No. 3.6e-07;
Matches 59; Conservative 0; Mismatches 17; Indels 0; Gaps 0;
Qy 1 GTGACAGACGGCTCGCCCAATTGAACCTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAA 60
Db 229 GTGAACCATGGAGGACCCCATGAACTTTCAGAGAAGTTGTTGATGATTCATAGATGAA 170
Qy 61 GTACCAATGGCTGTGA 76

Db 169 GTACCAATGGCTGTGA 154

RESULT 13

US-10-356-708-25/c

Sequence 25, Application US/10356708

Publication No. US20030157682A1

GENERAL INFORMATION:

APPLICANT: Large Scale Biology Corporation

TITLE OF INVENTION: MISMATCH ENDONUCLEASES AND METHODS OF USE THEREOF

FILE REFERENCE: P-LG 10100

CURRENT APPLICATION NUMBER: US/10/356,708

CURRENT FILING DATE: 2003-01-31

PRIOR FILING DATE: 2003-01-31

PRIOR FILING DATE: 2000-02-01

PRIOR FILING DATE: 2000-02-01

PRIOR FILING DATE: 2002-03-14

PRIOR FILING DATE: 2002-03-14

PRIOR FILING DATE: 2002-08-01

NUMBER OF SEQ ID NOS: 35

SOFTWARE: PatentIn version 3.1

SEQ ID NO 25

LENGTH: 769

TYPE: DNA

ORGANISM: Artificial Sequence derived from ToMV & TMV-U2

FEATURE:

OTHER INFORMATION: This sequence was derived by shuffling in accordance with the method of the present invention.

US-10-356-708-25

Query Match 64.2%; Score 48.8; DB 15; Length 769;

Best Local Similarity 77.6%; Pred. No. 3.6e-07;

Matches 59; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GTGACAGACGGCTCGCCAAATTTGAACCTCACTGAAAGGTTGTTGAGGAGTTTCGTGGATGAA 60

Db 229 GTGACAGATGGAGACCCATGGAACCTTTCAGAGAAGTTGTTGATGAGTTTCATAGATGAA 170

Qy 61 GTACCAATGGCTGTGA 76

Db 169 GTACCAATGGCTGTGA 154

RESULT 14

US-10-280-913A-25/c

Sequence 25, Application US/10280913A

Publication No. US20040110130A1

GENERAL INFORMATION:

APPLICANT: Large Scale Biology Corporation

TITLE OF INVENTION: POPULATION OF POLYNUCLEOTIDE SEQUENCE VARIANTS

FILE REFERENCE: LSBC-PLG5392-CIP

CURRENT APPLICATION NUMBER: US/10/280,913A

CURRENT FILING DATE: 2002-10-25

PRIOR APPLICATION NUMBER: 60/402,342

PRIOR FILING DATE: 2002-08-08

PRIOR APPLICATION NUMBER: 10/066,390

PRIOR FILING DATE: 2002-02-01

PRIOR APPLICATION NUMBER: 60/268,785

PRIOR FILING DATE: 2001-02-14

PRIOR APPLICATION NUMBER: 60/266,386

NUMBER OF SEQ ID NOS: 35

SOFTWARE: PatentIn version 3.1

SEQ ID NO 25

LENGTH: 769

TYPE: DNA

ORGANISM: Artificial Sequence

FEATURE:

OTHER INFORMATION: This construct was derived by GRAMMR shuffling in accordance with the method of the present invention.

US-10-280-913A-25

Query Match 64.2%; Score 48.8; DB 17; Length 769;

Best Local Similarity 77.6%; Pred. No. 3.6e-07;

Matches 59; Conservative 0; Mismatches 17; Indels 0; Gaps 0;

Qy 1 GTGACAGACGGCTCGCCAAATTTGAACCTCACTGAAAGGTTGTTGAGGAGTTTCGTGGATGAA 60

Db 229 GTGACAGATGGAGACCCATGGAACCTTTCAGAGAAGTTGTTGATGAGTTTCATAGATGAA 170

Qy 61 GTACCAATGGCTGTGA 76

Db 169 GTACCAATGGCTGTGA 154

Search completed: January 17, 2005, 21:46:50

Job time : 26.3891 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2005, 05:56:43 ; Search time 150.251 Seconds
(without alignments)
18431.958 Million cell updates/sec

Title: US-09-551-494-5_COPY_5430_5505
Perfect score: 76
Sequence: 1 gtgacagagcgtcgcaat.....tgaagtacaaatggtgtga 76

Scoring table: IDENTITY NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: gb_est1:*
2: gb_est2:*
3: gb_hic:*
4: gb_est3:*
5: gb_est4:*
6: gb_est5:*
7: gb_est6:*
8: gb_gest1:*
9: gb_gest2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
c 1	39.4	51.8	188	4	BM068137 KS08017C0
c 2	39.4	51.8	348	4	BM067518 KS08006E1
c 3	32.6	42.9	583	6	CB264749 41-E01466
4	31.6	41.6	954	8	B12288 T2M2-Sp6 TA
5	31	40.8	389	1	AV441961 AV441961
6	31	40.8	530	6	CA963974 CATI008D1
7	31	40.8	591	9	FR0020421 F.rubripe
8	31	40.8	656	5	B0636714 010F08 In
9	31	40.8	686	5	B0636451 052H05 In
10	31	40.8	1535	3	CNS0ADUO Arabidops
11	31	40.8	1579	3	CNS0ACVY Arabidops
c 12	30.2	39.7	711	5	EX845438 BX845438
c 13	30.2	39.7	800	5	B0915084 AGENCOURT
c 14	30	39.5	446	9	AL013342 F.rubripe
c 15	29.8	39.2	561	5	BX553578 BX553578
c 16	29.6	38.9	357	9	CG742490 ZMMBBc019
17	29.6	38.9	854	7	CO800453 AGENCOURT
c 18	29.4	38.7	732	1	AU305571 AU305571
19	28.8	37.9	388	5	B0816956 UAL1BPB12
20	28.8	37.9	398	9	CNS00STV Arabidops
c 21	28.8	37.9	460	7	CK095490 UAL1BPB12
c 22	28.8	37.9	498	8	AZ912881 RPCI-24-1
23	28.8	37.9	596	1	AJ769214 AJ769214
24	28.8	37.9	605	1	AJ774675 AJ774675

25	28.8	37.9	643	1	AJ767586
c 26	28.8	37.9	946	8	CC099545
c 27	28.6	37.6	514	8	BZ590961
c 28	28.6	37.6	516	8	BZ590922
c 29	28.6	37.6	522	8	BZ590931
c 30	28.6	37.6	527	8	BZ588716
c 31	28.6	37.6	529	8	BZ586702
c 32	28.6	37.6	548	7	CN947691
c 33	28.6	37.6	594	8	BZ586917
c 34	28.6	37.6	763	9	CC621362
c 35	28.6	37.6	877	9	CG305892
c 36	28.6	37.6	1003	8	BZ700933
c 37	28.2	37.1	385	9	CG742799
38	28.2	37.1	616	9	CC681582
39	28.2	37.1	635	9	CG392856
40	28.2	37.1	938	9	CL238374
41	28	36.8	244	9	CL226782
42	28	36.8	264	9	CG838137
43	28	36.8	276	9	CG743187
44	28	36.8	300	9	CG831760
45	28	36.8	547	6	CD836299

ALIGNMENTS

RESULT 1 BM068137/c 188 bp mRNA linear EST 11-SEP-2002
LOCUS KS08017C07 KS08 Capsicum annuum cDNA, mRNA sequence.
DEFINITION BM068137
ACCESSION BM068137
VERSION BM068137.1 GI:22788242
KEYWORDS EST.
SOURCE Capsicum annuum
ORGANISM Capsicum annuum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Capsicum.
REFERENCE 1 (bases 1 to 188)
AUTHORS Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S.,
Hur, C.-G. and Choi, D.
TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
annuum L.) and Sequence Analysis in Relation to Hypersensitive
Response Against Pathogen
JOURNAL Unpublished (2001)
COMMENT Contact: Doil Choi
Genome Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yusong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doil@ail.kribb.re.kr
High quality sequence stop: 188.

FEATURES
source
1..188
/organism="Capsicum annuum"
/mol_type="mRNA"
/cultivar="Hang Keun"
/db_xref="taxon:4072"
/tissue_type="anther"
/dev_stage="10 weeks after germination"
/clone_lib="KS08"
/note="Vector: pBluescript SK(-)"

ORIGIN
Query Match 51.8%; Score 39.4; DB 4; Length 188;
Best Local Similarity 71.2%; Pred.No. 0.01; Indels 0; Gaps 0;
Matches 52; Conservative 0; Mismatches 21; Indels 0; Gaps 0;
Qy 1 GTGACAGACGGCTCGCCAAATGAACCTCACTGAAAGGTTGTGAGGAGTTGTCGTGATGAA 60
|||||
Db 122 GTGTGCGAAGGAGGACCCGTTGAACCTTACAGAGCACTGTTGTGATGATTCATCGAATCA 63
|||||

```

QY 61 GTACCAATGGCTG 73
Db 62 GTTCCAATGGCTG 50

RESULT 2
BM067518/c
LOCUS BM067518 348 bp mRNA linear EST 11-SEP-2002
DEFINITION KS08006E10 KS08 Capsicum annuum cDNA, mRNA sequence.
ACCESSION BM067518
VERSION BM067518.1 GI:22787638
KEYWORDS EST.
SOURCE Capsicum annuum
ORGANISM Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Capsicum.
REFERENCE 1 (bases 1 to 348)
AUTHORS Lee,S., Kim,S.-Y., Chung,Y.-H., Shin,H.-J., Goh,S.-H., Pai,H.-S.,
Hur,C.-G. and Choi,D.
TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
annuum L.) and Sequence Analysis in Relation to Hypersensitive
Response Against Pathogen
JOURNAL Unpublished (2001)
COMMENT Contact: Doil Choi
Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yusong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doil@mail.kribb.re.kr
High quality sequence stop: 348.
FEATURES
    source
        location/Qualifiers
            1..348
                /organism="Capsicum annuum"
                /mol_type="mRNA"
                /cultivar="Hang Keun"
                /db_xref="taxon:4072"
                /tissue_type="anther"
                /dev_stage="10 weeks after germination"
                /clone_lib="KS08"
                /note="Vector: pBluescript SK(-)"

ORIGIN
Query Match 51.8%; Score 39.4; DB 4; Length 348;
Best Local Similarity 71.2%; Pred. No. 0.011;
Matches 52; Conservative 0; Mismatches 21; Indels 0; Gaps 0;

QY 1 GTGACAGCGCTCGCCAAATGAACTCACTGAAAGGTTGTTGAGGATTCGTGGATGAA 60
Db 294 GTGTCGGAAGAGGACCGGTGAACTTACAGACGAGTTGTTGATGAGTTTCATCGAATCA 235

QY 61 GTACCAATGGCTG 73
Db 234 GTTCCAATGGCTG 222

RESULT 3
CB264749
LOCUS CB264749 583 bp mRNA linear EST 06-NOV-2003
DEFINITION 41-E014660-035-002-B11-T7R MP1Z-ADIS-035 Arabidopsis thaliana cDNA
clone MP1Zp2000B11Q2 5-PRIME, mRNA sequence.
ACCESSION CB264749
VERSION CB264749.1 GI:32889522
KEYWORDS EST.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 583)
AUTHORS Schmid,K.J., Soerensen,T.R., Stracke,R., Torjek,O., Altmann,T.,
Mitchell-Olds,T. and Weisshaar,B.

TITLE Large-scale identification and analysis of genome-wide
single-nucleotide polymorphisms for mapping in Arabidopsis thaliana
Genome Res. 13 (6), 1250-1257 (2003)
JOURNAL 22683290
MEDLINE 12799357
PUBMED 12799357
COMMENT Contact: Weisshaar B
ADIS DNA core facility at MP1Z
Max-Planck-Institute for Plant Breeding Research
Carl-von-Linne Weg 10, 50829 Koeln, Germany
Fax: 00492215062851
Email: weisshaaem@piz-koeln.mpg.de
Insert Length: 583 Std Error: 0.00
Plate: 2 row: B column: 11
Seq primer: T7R; CTAATACGACTCACTATAGGA.
FEATURES
    source
        location/Qualifiers
            1..583
                /organism="Arabidopsis thaliana"
                /mol_type="mRNA"
                /cultivar="Achkarren-2 (AK-2)"
                /db_xref="GABI:594666"
                /db_xref="taxon:3702"
                /clone="MP1Zp2000B11Q2"
                /tissue_type="inflorescence"
                /lab_host="E. coli TOP10"
                /clone_lib="MP1Z-ADIS-035"
                /note="Vector: pSPORT1; Site 1: Sali; Site 2: NotI; cDNA
                library from Arabidopsis thaliana, accession Achkarren-2;
                inflorescences from flower buds to young siliques; library
                was made at the Max-Planck-Institute for Plant Breeding
                Research, Cologne, Germany; cloning sites Sali-NotI,
                primer sites and orientation:
                T7-Sali-CCACGCTCCG-5prime-cDNA-polyA-CC-NotI-SP6; Note:
                Sequencing granted in the context of the GABI Arabidopsis
                VerbundII: Genetic Diversity, 'Establishment of
                high-efficiency SNP-based mapping tools and development of
                methods for genome-wide mutation managed by
                Weisshaar; Sequence submission managed by
                RZPD/GABI-Primary database: http://gabi.rzpd.de This clone
                is available from RZPD; contact RZPD (clone@rzpd.de) for
                further information."

ORIGIN
Query Match 42.9%; Score 32.6; DB 6; Length 583;
Best Local Similarity 66.2%; Pred. No. 2.6;
Matches 47; Conservative 0; Mismatches 24; Indels 0; Gaps 0;

QY 6 AGACGCTCGCCAAATGAACTCACTGAAAGGTTGTTGAGGATTCGTGGATGAATACC 65
Db 381 AGATGCTTCTCCACTTGAATCATGATGAAGCTCTTGAGAAATTTGGAGACGAATCGC 440

QY 66 AATGGCTGTGA 76
Db 441 AATTGCTTCA 451

RESULT 4
B12288
LOCUS B12288 954 bp DNA linear GSS 14-MAY-1997
DEFINITION T2M2-Sp6 TAMU Arabidopsis thaliana genomic clone T2M2, genomic
survey sequence.
ACCESSION B12288
VERSION B12288.1 GI:2093409
KEYWORDS GSS.
SOURCE Arabidopsis thaliana (thale cress)
ORGANISM Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE 1 (bases 1 to 954)
AUTHORS Feng,J., Dewar,K., Buehler,E., Kim,C., Li,Y., Shinn,P., Sun,H. and
Ecker,J.
TITLE BAC End Sequences at ATGC
JOURNAL Unpublished (1997)

```

```

COMMENT
Other GSSs: T2M2-T7
Contact: Ecker J.
Arabidopsis thaliana Genome Center
University of Pennsylvania
Dept. of Biology, University of Pennsylvania, Philadelphia, PA
19104
Tel: 215-898-9384
Fax: 215-898-8780
Email: jecker@genome.bio.upenn.edu
Seq primer: Sp6
Class: BAC ends
High quality sequence start: 93
High quality sequence stop: 103.
FEATURES
    source
        1..954
            /organism="Arabidopsis thaliana"
            /mol_type="genomic DNA"
            /ecotype="Columbia"
            /db_xref="taxon:3702"
            /clone="T2M2"
            /sex="hermaphrodite"
            /clone_lib="TAMU"
            /notes="Vector: BelobACII; Site 1: HindIII; Site 2:
            HindIII; Produced by Rod Wing"
ORIGIN
Query Match 41.6%; Score 31.6; DB 8; Length 954;
Best Local Similarity 64.8%; Pred. No. 6.3;
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
Qy 6 AGACGGCTCCCAATGAACCTCACTGAAAGGTTGTGAGGAGTTCGTGGATGAAGTACC 65
Db 72 ANATGCTCTCCACTTGAATCATGATAAAGCTCTTGAGAGATTCGGAGACAATTCGC 131
Qy 66 AATGGCTGTGA 76
Db 132 AATGCTTTTA 142
RESULT 5
AV441961
LOCUS
DEFINITION
AV441961 Arabidopsis thaliana above-ground organ two to six-week
old Arabidopsis thaliana cDNA clone APD19f09_r 5', mRNA sequence.
ACCESSION
AV441961
VERSION
AV441961.1 GI:7612359
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 389)
Asamizu E., Nakamura Y., Sato S. and Tabata S.
A large scale analysis of cDNA in Arabidopsis thaliana: Generation
of 12,028 non-redundant expressed sequence tags from normalized and
size-selected cDNA libraries
DNA Res. 7 (3), 175-180 (2000)
20363093
10907847
PUBMED
COMMENT
Contact: Erika Asamizu
The First Laboratory for Plant Gene Research
Kazusa DNA Research Institute
Yana 1532-3, Kisarazu, Chiba 292-0812, Japan
Email: asamizu@kazusa.or.jp, URL: http://www.kazusa.or.jp/en/plant/.
FEATURES
    source
        1..389
            /organism="Arabidopsis thaliana"
            /mol_type="mRNA"
            /ecotype="Columbia"
            /db_xref="taxon:3702"
            /clone="APD19f09_r"
            /tissue_type="aboveground organs"

```

```

/dev_stage="two to six-week old"
/clone_lib="Arabidopsis thaliana above-ground organ two to
six-week old"
/notes="Vector: pBluescriptII SK-; Site_1: EcoRI; Site_2:
XhoI"
ORIGIN
Query Match 40.8%; Score 31; DB 1; Length 389;
Best Local Similarity 64.8%; Pred. No. 8.5;
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
Qy 6 AGACGGCTCCCAATGAACCTCACTGAAAGGTTGTGAGGAGTTCGTGGATGAAGTACC 65
Db 62 AGATGCTTCTCCACTTGAATCATGATAAAGCTCTTGAGAGATTCGGAGACAATTCGC 121
Qy 66 AATGGCTGTGA 76
Db 122 AATGCTTTTA 132
RESULT 6
CA963974
LOCUS
DEFINITION
CA963974 530 bp mRNA linear EST 03-JAN-2003
CATION08D10AF Infected Arabidopsis Leaf Arabidopsis thaliana cDNA,
mRNA sequence.
ACCESSION
CA963974
VERSION
CA963974.1 GI:27490531
KEYWORDS
EST.
SOURCE
Arabidopsis thaliana (thale cress)
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 530)
Lundgaard M., Emmersen J., Nielsen K.L., Wilson I., Somerville S.
EST sequencing of Erysiphe cichoracearum infected Arabidopsis
plants
Unpublished (2002)
Contact: Karen G. Welinder
Institut for bioteknologi
Aalborg Universitet
Sohnsgaardsholmsvej 49, 9000 Aalborg, Denmark
Tel: +45 96358467
Fax: +45 98141808
Email: kgw@bio.auc.dk.
FEATURES
    source
        1..530
            /organism="Arabidopsis thaliana"
            /mol_type="mRNA"
            /ecotype="Columbia"
            /db_xref="taxon:3702"
            /dev_stage="Plant 3 weeks old, three days post infection"
            /clone_lib="Infected Arabidopsis Leaf"
            /notes="Organ: Leaf; Vector: pBluescript; Mixed cDNA
            library of Arabidopsis and E. cichoracearum infected leaf
            from three weeks old Arabidopsis plants. Plants were
            harvested 3 days after infection and mRNA oligo dt
            selected."
ORIGIN
Query Match 40.8%; Score 31; DB 6; Length 530;
Best Local Similarity 64.8%; Pred. No. 9;
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
Qy 6 AGACGGCTCCCAATGAACCTCACTGAAAGGTTGTGAGGAGTTCGTGGATGAAGTACC 65
Db 215 AGATGCTTCTCCACTTGAATCATGATAAAGCTCTTGAGAGATTCGGAGACAATTCGC 274
Qy 66 AATGGCTGTGA 76
Db 275 AATGCTTTTA 285

```

```

RESULT 7
FR0020421
LOCUS      F0020421
DEFINITION F.rubripes GSS sequence, clone 041P11dB11, genomic survey sequence.
ACCESSION  AL013304
VERSION     AL013304.1 GI:2679672
KEYWORDS   GSS; Genomic survey sequence.
SOURCE     Takifugu rubripes (Fugu rubripes)
ORGANISM   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
            Acanthomorpha; Acanthopterygii; Percomorpha; Tetraodontiformes;
            Tetraodontidae; Tetraodontidae; Takifugu.
REFERENCE  1
AUTHORS    Elgar,G., Clark,M.S., Meek,S., Smith,S., Warner,S., Williams,G. and
            Bouchireb,N., Cottage,A., Yeo,G.S., Umrana,Y., Williams,G. and
            Brenner,S.
TITLE      Generation and analysis of 25 Mb of genomic DNA from the pufferfish
JOURNAL    Fugu rubripes by sequence scanning
MEDLINE     Genome Res. 9 (10), 960-971 (1999)
PUBMED      99455097
REFERENCE   2 (bases 1 to 591)
AUTHORS    Elgar,G., Clark,M., Smith,S., Meek,S., Warner,S., Umrana,Y.,
            Williams,G. and Brenner,S.
TITLE      Direct Submission
JOURNAL     Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource
COMMENT     Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@gmp.mrc.ac.uk
            Vector: pBluescript II KS
            V type: phagemid
            PRIMER: KS
DESCR:      One pass dye-terminator sequencing of cosmid cloned genomic
            sequence.
FEATURES   source
            1..591
            /organism="Takifugu rubripes"
            /mol_type="genomic DNA"
            /db_xref="taxon:31033"
            /clones="041P11dB11"
            /clone_lib="cosmid 041P11"
ORIGIN
Query Match      40.8%; Score 31; DB 9; Length 591;
Best Local Similarity 64.2%; Pred. No. 9.2;
Matches 43; Conservative 0; Mismatches 24; Indels 0; Gaps 0;
Qy 10 GGCTCGCAATTGAACCTCACTGAAAGCTTCTTGAGGAGTTCGTGGATGAGTACCAATG 69
Db 113 GGGCTGCGAGTTGAACCTGGAANATGTTGACGATCTTGGGGACGAACTGTAAAG 172
Qy 70 GCTGTGA 76
Db 173 GTTGCTGA 179

RESULT 8
BU636714
LOCUS      BU636714
DEFINITION 010F08 Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA
sequence.
ACCESSION  BU636714
VERSION     BU636714.1 GI:23303969
KEYWORDS   EST.
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 656)
AUTHORS    Lundsgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S.
            and Wellinder,K.G.
TITLE      EST sequencing of Erysiphe cichoracearum infected Arabidopsis
JOURNAL     Unpublished (2002)
COMMENT     Contact: Karen G. Wellinder
            Institut for bioteknologi
            Aalborg Universitet
            Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
            Tel: +45 96358467
            Fax: +45 98141808
            Email: kgw@bio.auc.dk.
            Location/Qualifiers
            1..656
            /organism="Arabidopsis thaliana"
            /mol_type="mRNA"
            /db_xref="taxon:31033"
            /dev_stage="Plant 3 weeks old, three days post infection"
            /clone_lib="Infected Arabidopsis Leaf"
            /note="Organ: Leaf; Vector: pBluescript; Mixed cDNA
            from three weeks old Arabidopsis plants. Plants were
            harvested 3 days after infection and mRNA oligo dt
            selected."
ORIGIN
Query Match      40.8%; Score 31; DB 5; Length 656;
Best Local Similarity 64.8%; Pred. No. 9.4;
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
Qy 6 AGAGCGCTCCCAATTGAACCTCACTGAAAGTGTGTTGAGGAGTTCGTGGATGAAGTACC 65
Db 363 AGATGCTTCTCCACTTGAATCATGTAAGCTCTTGAGAGATTCGGAGACCAATCGC 422
Qy 66 AATGGCTGTGCA 76
Db 423 AATGCTTTTGA 433

RESULT 9
BU636451
LOCUS      BU636451
DEFINITION 052H05 Infected Arabidopsis Leaf Arabidopsis thaliana cDNA, mRNA
sequence.
ACCESSION  BU636451
VERSION     BU636451.1 GI:23303706
KEYWORDS   EST.
SOURCE     Arabidopsis thaliana (thale cress)
ORGANISM   Arabidopsis thaliana
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
REFERENCE   1 (bases 1 to 686)
AUTHORS    Lundsgaard,M., Emmersen,J., Nielsen,K.L., Wilson,I., Somerville,S.
            and Wellinder,K.G.
TITLE      EST sequencing of Erysiphe cichoracearum infected Arabidopsis
JOURNAL     Unpublished (2002)
COMMENT     Contact: Karen G. Wellinder
            Institut for bioteknologi
            Aalborg Universitet
            Sohngaardsholmsvej 49, 9000 Aalborg, Denmark
            Tel: +45 96358467
            Fax: +45 98141808
            Email: kgw@bio.auc.dk.
            Location/Qualifiers
            1..686
            /organism="Arabidopsis thaliana"
            /mol_type="mRNA"
            /db_xref="taxon:31033"
            /dev_stage="Plant 3 weeks old, three days post infection"
            /clone_lib="Infected Arabidopsis Leaf"
            /note="Organ: Leaf; Vector: pBluescript; Mixed cDNA
            from three weeks old Arabidopsis plants. Plants were
            harvested 3 days after infection and mRNA oligo dt
            selected."
ORIGIN
Query Match      40.8%; Score 31; DB 5; Length 686;
Best Local Similarity 64.8%; Pred. No. 9.4;
Matches 46; Conservative 0; Mismatches 25; Indels 0; Gaps 0;
Qy 6 AGAGCGCTCCCAATTGAACCTCACTGAAAGTGTGTTGAGGAGTTCGTGGATGAAGTACC 65
Db 363 AGATGCTTCTCCACTTGAATCATGTAAGCTCTTGAGAGATTCGGAGACCAATCGC 422
Qy 66 AATGGCTGTGCA 76
Db 423 AATGCTTTTGA 433

```

Qy 6 AGACGGCTCGCCAAATTGAACTCACCTGAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACC 65

```

Db      368 AGATGCTTCTCCACTTGAGATCATGGATAGAGCTCTTGAGAGATTGGAGACCAATCGC 427
QY      66 AATGCTGTGA 76
      |||||
Db      428 AATGCTTTTA 438

RESULT 12
BX845438/c
LOCUS   BX845438
DEFINITION
  BX845438 NICHD_XGC_OO1 Xenopus laevis cDNA clone IMAGE998P2114225 ;
  IMAGE:6641733 5', mRNA sequence.
ACCESSION
  BX845438
VERSION
  BX845438.1 GI:39735223
KEYWORDS
  EST.
SOURCE
  Xenopus laevis (African clawed frog)
ORGANISM
  Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
  Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
  Xenopodinae; Xenopus; Xenopus.

REFERENCE
  Heil, O., Neubert, P., Peters, M., Radelof, U., Schneider, D.,
  Schroth, A., Korn, B. and Landgrebe, J.
  Xenopus laevis UniGene Set 1 (RZPDLib No. 988)
  Unpublished (2003)
  Contact: Ina Rolfs
  RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
  Im Neuenheimer Feld 580, D-69120 Heidelberg, Germany
  RZPD; IMAGE998P2114225.
  RZPDLib; I.M.A.G.E. cDNA Clone Collection (amp- resistant) (RZPDLib
  No. 998) http://www.rzpd.de/cgi-
  bin/products/showLib.pl.cgi/response?libNo=998 RZPDLib; Xenopus
  laevis UniGene Set 1 (RZPDLib No. 988) http://www.rzpd.de/cgi-
  bin/products/showLib.pl.cgi/response?libNo=988 Contact: Ina Rolfs
  RZPD Deutsches Ressourcenzentrum fuer Genomforschung GmbH
  Heubnerweg 6, D-14059 Berlin, Germany
  Tel: +49 30 32639 101
  Fax: +49 30 32639 111
  www.rzpd.de
  This clone is available royalty-free from RZPD;
  contact RZPD (clone@rzpd.de) for further information. Seq primer:
  SP6, 5' ATTAGGTGACATATAG 3'.
FEATURES
  source
    1..711
    /organism="Xenopus laevis"
    /mol_type="mRNA"
    /db_xref="taxon:8355"
    /clone="IMAGE998P2114225 ; IMAGE:6641733"
    /tissue_type="oocytes"
    /lab_host="DH10B (phage-resistant)"
    /clone_lib="NICHD_XGC_OO1"
    /notes="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;
    Cloned unidirectionally. Primer: Oligo dT. Average insert
    size 2.2 kb. Constructed by Life Technologies."
ORIGIN
  Query Match 39.7%; Score 30.2; DB 5; Length 711;
  Best Local Similarity 69.5%; Pred. No. 18;
  Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY      16 CCAATTGAACCTACTGAAAAGGTTGTTGAGAGTTCTGTGATGAAGTACCAATGGCTGT 74
      |||||
Db      210 CCAACAGAACTCACTGAAGAGAGCTGGGGATGTGGGGATGGAGCACCAGTGGTTT 152

RESULT 13
BU915084/c
LOCUS   BU915084
DEFINITION
  AGENCOURT 10494416 NICHD_XGC_OO1 Xenopus laevis cDNA clone
  IMAGE:6641710 5', mRNA sequence.
ACCESSION
  BU915084
VERSION
  BU915084.1 GI:24096998
KEYWORDS
  EST.

Xenopus laevis (African clawed frog)
Xenopus laevis
Eukaryota; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidae;
Xenopodinae; Xenopus; Xenopus.
1 (bases 1 to 800)
NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
Unpublished (1997)
Contact: Robert Strausberg, Ph.D.
Email: cgapbs-remail.nih.gov
Tissue Procurement: Martha Rabbert, Steven L. Klein, Ph.D.
CDNA Library Preparation: Life Technologies, Inc.
CDNA Library Arrayed by: The I.M.A.G.E. Consortium (LLNL)
DNA sequencing by: Agencourt Bioscience Corporation
Clone distribution: NCI-CGAP clone distribution information can be
found through the I.M.A.G.E. Consortium/LLNL at:
http://image.llnl.gov
Plate: L1AM14225 row: O column: 22
High quality sequence stop: 590.
Location/Qualifiers
  1..800
  /organism="Xenopus laevis"
  /mol_type="mRNA"
  /db_xref="taxon:8355"
  /clone="IMAGE:6641710"
  /tissue_type="oocytes"
  /lab_host="DH10B (phage-resistant)"
  /clone_lib="NICHD_XGC_OO1"
  /notes="Vector: pCMV-SPORT6; Site 1: NotI; Site 2: SalI;
  Cloned unidirectionally. Primer: Oligo dT. Average insert
  size 2.2 kb. Constructed by Life Technologies."
ORIGIN
  Query Match 39.7%; Score 30.2; DB 5; Length 800;
  Best Local Similarity 69.5%; Pred. No. 18;
  Matches 41; Conservative 0; Mismatches 18; Indels 0; Gaps 0;

QY      16 CCAATTGAACCTACTGAAAAGGTTGTTGAGAGTTCTGTGATGAAGTACCAATGGCTGT 74
      |||||
Db      210 CCAACAGAACTCACTGAAGAGAGCTGGGGATGTGGGGATGGAGCACCAGTGGTTT 152

RESULT 14
FR0020459/c
LOCUS   FR0020459
DEFINITION
  F. rubripes GSS sequence, clone 041PlidF7, genomic survey sequence.
ACCESSION
  AL013342
VERSION
  AL013342.1 GI:2679710
KEYWORDS
  GSS; genome survey sequence.
  Takifugu rubripes (Pugu rubripes)
SOURCE
  Takifugu rubripes
ORGANISM
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
  Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;
  Tetraodontidae; Tetraodontidae; Takifugu.
1
Elgar, G., Clark, M.S., Meek, S., Smith, S., Warner, S., Edwards, Y.J.,
Bouchireb, N., Cottage, A., Yeo, G.S., Umrana, Y., Williams, G. and
Brenner, S.
Generation and analysis of 25 Mb of genomic DNA from the pufferfish
Fugu rubripes by sequence scanning
Genome Res. 9 (10), 960-971 (1999)
99455097
10523524
2 (bases 1 to 446)
Elgar, G., Clark, M., Smith, S., Meek, S., Warner, S., Umrana, Y.,
Williams, G. and Brenner, S.
Direct Submission
Submitted (08-DEC-1997) MRC Human Genome Mapping Project Resource
Centre Hinxton, Cambridge, CB10 1SB. Email: biohelp@hgmrc.ac.uk
Vector: pBluescript II KS

```

V type: phagemid
 PRIMER: KS
 DESCR: One pass dye-terminator sequencing of cosmid cloned genomic sequence.
 FEATURES
 source
 Location/Qualifiers
 1..446
 /organism="Takifugu rubripes"
 /mol_type="genomic DNA"
 /db_xref="taxon:31033"
 /clone="041P11DF"
 /clone_lib="cosmid 041P11"
 ORIGIN
 Query Match 39.5%; Score 30; DB 9; Length 446;
 Best Local Similarity 64.6%; Pred. No. 19;
 Matches 42; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
 QY 10 GGCTCGCAATTGCACTCACTGAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACCAATG 69
 |||||
 Db 408 GGCTCGGAGTTGAACTCTTGGCAGATGTTGTTCAGCATCTTGGGGACGAANANNAAG 349
 |||||
 QY 70 GCTGT 74
 |||||
 Db 348 GTTGT 344

RESULT 15
 BX553578/c
 LOCUS
 DEFINITION BX553578 Glossina morsitans morsitans adult infected gut Glossina morsitans morsitans cDNA clone Tse12c12_q1c, mRNA sequence.
 ACCESSION BX553578
 VERSION BX553578.1 GI:33377725
 KEYWORDS EST.
 SOURCE Glossina morsitans morsitans
 ORGANISM Glossina morsitans morsitans
 Eukaryota; Metazoa; Arthropoda; Hexapoda; Insecta; Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha; Hippoboscidae; Glossinidae; Glossina.
 REFERENCE 1 (bases 1 to 561)
 AUTHORS Lehane, M.J., Aksoy, S., Gibson, W., Kerhornou, A., Berriman, M., Hamilton, J., Soares, M.B., Bonaldo, M.F., Lehane, S. and Hall, N.
 TITLE Adult midgut expressed sequence tags from the tsetse fly Glossina morsitans morsitans and expression analysis of putative immune response genes
 JOURNAL Genome Biol. 4 (10), R63 (2003)
 MEDLINE 22881942
 PUBMED 14519198
 COMMENT Contact: Hall N
 Pathogen Sequencing Unit
 The Sanger Institute The Wellcome Trust Genome Campus
 Hinxton, Cambridge, CB10 1SA, UK
 Request for clones, please contact: Mike Lehane
 Prof. M.J. Lehane
 School of Biological Sciences,
 University of Wales,
 Bangor LL57 2UW
 All clones with suffix q1c are reverse primer reads starting at 5' end of the cDNA all p1c reads are from the 3' end.
 FEATURES
 source
 Location/Qualifiers
 1..561
 /organism="Glossina morsitans morsitans"
 /mol_type="mRNA"
 /sub_species="morsitans"
 /db_xref="taxon:37546"
 /clone="Tse12c12_q1c"
 /tissue_type="adult infected gut"
 /clone_lib="Glossina morsitans morsitans adult infected gut"
 /note="country: Zimbabwe; EST from adult gut infected with T.brucei"

ORIGIN
 Query Match 39.2%; Score 29.8; DB 5; Length 561;
 Best Local Similarity 66.2%; Pred. No. 23;
 Matches 43; Conservative 0; Mismatches 22; Indels 0; Gaps 0;
 QY 5 CAGACGGCTCGCCCAATTGAACCTCACTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTAC 64
 |||||
 Db 341 CTGAAGAATCGGCCAATTTCTTGACTGTTATGTTGTTAGTAGATTAGTGGATGATGATAC 282
 |||||
 QY 65 CAATG 69
 |||||
 Db 281 TAACG 277

Search completed: January 17, 2005, 16:28:48
 Job time : 154.251 secs

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2005, 21:55:52 ; Search time 17605.5 Seconds
(without alignments)
17070.053 Million cell updates/sec

Title: US-09-551-494-5
Perfect score: 6355
Sequence: 1 gatgtttaatagtttcga.....taaccgcgtagggccca 6355

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4526729 seqs, 23644849745 residues

Total number of hits satisfying chosen parameters: 9053458

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database :

GenEmbl.*

1: gb_ba.*

2: gb_hcg.*

3: gb_in.*

4: gb_om.*

5: gb_ov.*

6: gb_pat.*

7: gb_ph.*

8: gb_pl.*

9: gb_pr.*

10: gb_ro.*

11: gb_sts.*

12: gb_sy.*

13: gb_un.*

14: gb_vl.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6355	100.0	6355	6	AX040177 Sequence
2	6355	100.0	6355	6	AX795380 Sequence
3	6355	100.0	6355	14	M34077 Tobacco mil
4	5951.6	93.7	6356	14	AB078435 Tobacco m
5	2631	41.4	6384	6	E07987 Tobacco mos
6	2631	41.4	6384	14	X02144 Tobacco mos
7	2631	41.4	6385	14	AB083196 Tobacco mo
8	2626.2	41.3	6384	14	AJ132845 tobacco mo
9	2623.8	41.3	6383	14	AJ243571 tobacco m
10	2621.6	41.3	6384	14	AF155507 Tobacco m
11	2620.6	41.2	6383	14	AF332868 Tobacco mo
12	2617.4	41.2	6383	14	292909 Tobacco mos
13	2612.6	41.1	6383	14	AJ417701 Tomato mo
14	2529.6	39.8	6357	14	AB069853 Pepper mi
15	2528.4	39.8	6357	14	M81413 Nucleotide
16	2528	39.8	6357	14	AB000709 Pepper mi
17	2496.4	39.3	6357	14	AJ308228 Pepper Mi
18	2489.2	39.2	6395	14	D63809 Tobacco mos
19	2479.2	39.0	6395	14	AF395129 Tobacco m

20	2472.8	38.9	6395	14	AF395128	Tobacco m
21	2469.6	38.9	6395	14	AF273221	Tobacco m
22	2469.6	38.9	6395	14	AF395127	Tobacco m
23	2468	38.8	6395	6	AR173320	Sequence
24	2468	38.8	6395	6	AX098414	Sequence
25	2466.4	38.8	6395	6	AR271575	Sequence
26	2466.4	38.8	6395	6	AX040174	Sequence
27	2466.4	38.8	6395	14	TOTMV4	V01408 Tobacco mos
28	2461.6	38.7	6398	14	TOTMV5	V01409 Tobacco mos
29	2460	38.7	6395	14	TMVCG	X68110 Tobacco mos
30	2455.4	38.6	6439	6	AR173321	Sequence
31	2455.4	38.6	6439	6	AX098415	Sequence
32	2453.6	38.6	6395	14	TMV11933	AJ011933 Tobacco m
33	2452.2	38.6	6475	6	AR173323	Sequence
34	2452.2	38.6	6475	6	AX098417	Sequence
35	2452	38.6	6395	6	AX194392	Sequence
36	2452	38.6	6395	14	AF165190	AF165190 Tobacco m
37	2449.8	38.5	6446	6	AR173324	Sequence
38	2449.8	38.5	6446	6	AX098418	Sequence
39	2431.2	38.3	6425	6	AR173322	Sequence
40	2431.2	38.3	6425	6	AX098416	Sequence
41	2370.2	37.3	6524	14	AB089381	AB089381 Paprika m
42	2327.4	36.6	6506	14	MTVNGHYPER	L11665 Obuda peppe
43	2325.8	36.6	6507	14	MTVGRNA	D13438 Obuda peppe
44	2291.4	36.1	11641	6	AX466963	AX466963 Sequence
45	2286.6	36.0	7685	6	BD263877	BD263877 Viral exp

ALIGNMENTS

RESULT 1
AX040177
LOCUS AX040177 6355 bp DNA linear PAT 18-NOV-2000
DEFINITION Sequence 5 from Patent WO0063397.
ACCESSION AX040177
VERSION AX040177.1 GI:11230127
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Meulewaeter, P., Cornelisse, Jacobs J., van Eldik, G. and Metzlaiff, M.
TITLE Methods and means for delivering inhibitory rna to plants and applications thereof
JOURNAL Patent: WO 0063397-A 5 26-OCT-2000;
Aventis CropScience N.V. (BE)
FEATURES Location/Qualifiers
source 1..6355
/organism="synthetic construct"
/mol_type="unassigned DNA"
/db_xref="taxon:32630"
/note="cDNA copy of the nucleotide sequence of the genome of TMV-U2"

ORIGIN	Query Match	100.0%	Score 6355	DB 6	Length 6355
Best Local Similarity	100.0%	Pred. No. 0			
Matches 6355	Conservative 0	Mismatches 0	Indels 0	Gaps 0	
Qy	1	GATGTTTTTATAGCTTTTCGACACACAACTATAAACAACAAACACATATTACAACACACA	60		
Db	1	GATGTTTTTATAGCTTTTCGACACACAACTATAAACAACAAACACATATTACAACACACA	60		
Qy	61	AACAACAACAAATGGCACACATACATCTATAATTAGCAACGCCCTTCTTGAAGCGTGAG	120		
Db	61	AACAACAACAAATGGCACACATACATCTATAATTAGCAACGCCCTTCTTGAAGCGTGAG	120		
Qy	121	TGGTAAAAAACACTCTCGTTAATGACCTTGCAGAGCGCGCATGTACGATACGGCGGTGA	180		
Db	121	TGGTAAAAAACACTCTCGTTAATGACCTTGCAGAGCGCGCATGTACGATACGGCGGTGA	180		
Qy	181	AGATTTTAAAGCCCGGACCGTAGACCAAGGTCACTTTTCCAAAACACTATTAGCGGAGA	240		

Db 181 AGAATTTAAAGCCCGCGACCGGTAGACCAAGGTCACATTTTCCAAACATATTAGCGAAG 240
Qy 241 GCAACGCTTCTAGTCTCCAAACGCGTACCCGAGTTCCAGATTACCTTTTATAAATCTCA 300
Db 241 GCAACGCTTCTAGTCTCCAAACGCGTACCCGAGTTCCAGATTACCTTTTATAAATCTCA 300
Qy 301 AATGCGGTACACAGTTTGGCTGGAGTTTGAGAGCAATTAGAATTTGGAATATCTGATGCT 360
Db 301 AATGCGGTACACAGTTTGGCTGGAGTTTGAGAGCAATTAGAATTTGGAATATCTGATGCT 360
Qy 361 ACAAGTTCCTTAGATCCGACATATGATATAGTGGGAATTTGAGAGCAATTTGTT 420
Db 361 ACAAGTTCCTTAGATCCGACATATGATATAGTGGGAATTTGAGAGCAATTTGTT 420
Qy 421 CAAAGGACGGGATTAAGTGCATTTGCTGATGCGCCCAATCTGGACATACAGAGATATAATGAG 480
Db 421 CAAAGGACGGGATTAAGTGCATTTGCTGATGCGCCCAATCTGGACATACAGAGATATAATGAG 480
Qy 481 GCACGAAGGACAAAGGACTCAATTTGAGATGATATTTGTCGAGTTGTCGTTCTAACAA 540
Db 481 GCACGAAGGACAAAGGACTCAATTTGAGATGATATTTGTCGAGTTGTCGTTCTAACAA 540
Qy 541 GGTAAATTCCTGAGTTTCAAAGGGAGGCTTTTAAACAGGTATGCAAGGCTCCCAACGAGT 600
Db 541 GGTAAATTCCTGAGTTTCAAAGGGAGGCTTTTAAACAGGTATGCAAGGCTCCCAACGAGT 600
Qy 601 CTGCTGCTTAAACCTTTTCAAGATGTCGAATACATCCGCGAGAGATAGTGGTAGAAG 660
Db 601 CTGCTGCTTAAACCTTTTCAAGATGTCGAATACATCCGCGAGAGATAGTGGTAGAAG 660
Qy 661 ATACGCTGTGCTGCGACAGTTTGTATGATATTTCTGTCGATGATGATGAGCTGGTT 720
Db 661 ATACGCTGTGCTGCGACAGTTTGTATGATATTTCTGTCGATGATGATGAGCTGGTT 720
Qy 721 AATATCTAAGAAATATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 721 AATATCTAAGAAATATACATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Qy 781 AGACGAGCGAGGTAGCTTAATGAAATAGGCGCAACTTTTCAAAAGAGAGGTGATGA 840
Db 781 AGACGAGCGAGGTAGCTTAATGAAATAGGCGCAACTTTTCAAAAGAGAGGTGATGA 840
Qy 841 TGTTCCTTTTCTTCTGCTGATGAAAGTACCTTTTAAATATAGTCAATTAACAAATAT 900
Db 841 TGTTCCTTTTCTTCTGCTGATGAAAGTACCTTTTAAATATAGTCAATTAACAAATAT 900
Qy 901 CTTGCAATATGATGATTAATCTTACTTTCTGCTTCTAGTAGAATAGTTTACTTTAAGGA 960
Db 901 CTTGCAATATGATGATTAATCTTACTTTCTGCTTCTAGTAGAATAGTTTACTTTAAGGA 960
Qy 961 ATTTTATGCTACTAGGCTTAATATCTTGGTTTGTAAATTTTAAAGTAGATACCTATAT 1020
Db 961 ATTTTATGCTACTAGGCTTAATATCTTGGTTTGTAAATTTTAAAGTAGATACCTATAT 1020
Qy 1021 TCTGTACAAGAGTGTAGACAAGTAGGCTGTAGTAGTATGATGATGATGATGATGATGATGAT 1080
Db 1021 TCTGTACAAGAGTGTAGACAAGTAGGCTGTAGTAGTATGATGATGATGATGATGATGATGAT 1080
Qy 1081 AGACGCTTTCCTTACAAGAAACCTTTGGCCATGTTTCAACATGAAAGAGCAATCTTTAG 1140
Db 1081 AGACGCTTTCCTTACAAGAAACCTTTGGCCATGTTTCAACATGAAAGAGCAATCTTTAG 1140
Qy 1141 AGACGCGCTTCGTTTAACTTTTGGTTCCCTAGATGAGGACATGGTATAGTACCGCT 1200
Db 1141 AGACGCGCTTCGTTTAACTTTTGGTTCCCTAGATGAGGACATGGTATAGTACCGCT 1200
Qy 1201 GTTGTAGGGTTCATTAACGAGCAAAAGATGACAAGAGGTGAGTCAATTTGTTAATCGTGA 1260
Db 1201 GTTGTAGGGTTCATTAACGAGCAAAAGATGACAAGAGGTGAGTCAATTTGTTAATCGTGA 1260
Qy 1261 CTTGCTTTTACAGTGTCTTAATCATATCAGAAATATCAAGGCAAAAGCGTTTAACTTACCA 1320

Db 1261 CTTGCTTTTACAGTGTCTTAATCATATCAGAAATATCAAGCCAAAGGTTAACTTACCA 1320
Qy 1321 GAAAGTATATCTTTCTGTTGAGTCTATTAAGATCCCGGTGATTAATCAATGTTTACTTGC 1380
Db 1321 GAAAGTATATCTTTCTGTTGAGTCTATTAAGATCCCGGTGATTAATCAATGTTTACTTGC 1380
Qy 1381 TAGGTCTGAATGGGATGTAGATAAAGCAATTTCTTCAACCCCTTGTCAATGACTTTCTTCTT 1440
Db 1381 TAGGTCTGAATGGGATGTAGATAAAGCAATTTCTTCAACCCCTTGTCAATGACTTTCTTCTT 1440
Qy 1441 GCAGACTAAGCTGCTGCTTCAAGACGATATAGTAAATGGGAAAGTTTTCGTTCTTGGGA 1500
Db 1441 GCAGACTAAGCTGCTGCTTCAAGACGATATAGTAAATGGGAAAGTTTTCGTTCTTGGGA 1500
Qy 1501 TAAGACCACTTCTGAACTTATTTGGGATGAGTGGGCAAAATTTTTCGGAACGTTTTCCTC 1560
Db 1501 TAAGACCACTTCTGAACTTATTTGGGATGAGTGGGCAAAATTTTTCGGAACGTTTTCCTC 1560
Qy 1561 CACTATCAAAGAGAGATTGGTGAAGCAAAATTTCTGATGTTAAGTGAAGATGCTCTGAA 1620
Db 1561 CACTATCAAAGAGAGATTGGTGAAGCAAAATTTCTGATGTTAAGTGAAGATGCTCTGAA 1620
Qy 1621 GATCAAGATCCAGATCTGATGTTACATGCAAGAGACAGGTTTCTAGTCTGAAATACCAA 1680
Db 1621 GATCAAGATCCAGATCTGATGTTACATGCAAGAGACAGGTTTCTAGTCTGAAATACCAA 1680
Qy 1681 GTCTGAGAGTTACCGCATCTAGATATCAAGAGAGGCTTAGAAGAGCTGAGCAAAATGTA 1740
Db 1681 GTCTGAGAGTTACCGCATCTAGATATCAAGAGAGGCTTAGAAGAGCTGAGCAAAATGTA 1740
Qy 1741 CGACGCGTTATCAGAAATTTCTTAAAGGCTGCTGATAATTTTCGATATCCGGAAGTT 1800
Db 1741 CGACGCGTTATCAGAAATTTCTTAAAGGCTGCTGATAATTTTCGATATCCGGAAGTT 1800
Qy 1801 CAAAGACATGTCAGAGCTTTAGATGTTTCTGATGTTGCGAGCAAGGATGATGATGATGAT 1860
Db 1801 CAAAGACATGTCAGAGCTTTAGATGTTTCTGATGTTGCGAGCAAGGATGATGATGATGAT 1860
Qy 1861 AGTGGCGCGAGATPAGAGCGGTTTAACTCTTACTTTTGTATGAAAGGAGGATGATGAT 1920
Db 1861 AGTGGCGCGAGATPAGAGCGGTTTAACTCTTACTTTTGTATGAAAGGAGGATGATGAT 1920
Qy 1921 GGTAAAGGCTTTTAAAGCAACGCGCTTCTGAGCGGTTGATGTTTGAACCCGACATCCGA 1980
Db 1921 GGTAAAGGCTTTTAAAGCAACGCGCTTCTGAGCGGTTGATGTTTGAACCCGACATCCGA 1980
Qy 1981 AGAGGTGAACGTAATAAATTTTCTATTTCTGAGAAAGGAGATGTCCTGTGTGTCAGA 2040
Db 1981 AGAGGTGAACGTAATAAATTTTCTATTTCTGAGAAAGGAGATGTCCTGTGTGTCAGA 2040
Qy 2041 AAGTCAATGTTTGAAGATGCTAACTTAGAGCAACGAGGATGAGTCCCTCAACGATTT 2100
Db 2041 AAGTCAATGTTTGAAGATGCTAACTTAGAGCAACGAGGATGAGTCCCTCAACGATTT 2100
Qy 2101 CCATAAGGCTTCGTTGAGTGTGATTAACAAGCAAAATGGCATCGGTTGTCTACATCGG 2160
Db 2101 CCATAAGGCTTCGTTGAGTGTGATTAACAAGCAAAATGGCATCGGTTGTCTACATCGG 2160
Qy 2161 CTCACCTCAAAGTTCAACAAATGAAGAACTATGTCGACAGTTTGGCAGCTTCTGCTCCGC 2220
Db 2161 CTCACCTCAAAGTTCAACAAATGAAGAACTATGTCGACAGTTTGGCAGCTTCTGCTCCGC 2220
Qy 2221 CACTGTATCAAAATCTATGCAAGTCACTAAAGGATGAGTCCGTTATGATTTCTGATCCAG 2280
Db 2221 CACTGTATCAAAATCTATGCAAGTCACTAAAGGATGAGTCCGTTATGATTTCTGATCCAG 2280
Qy 2281 GGAGAAAGTGGTGTGGGATGTCACCTTGAAGAGTGGCTTCTCAAAACCTTGGCGCAA 2340
Db 2281 GGAGAAAGTGGTGTGGGATGTCACCTTGAAGAGTGGCTTCTCAAAACCTTGGCGCAA 2340
Qy 2341 AGGTCAATTCATGGGAGTTGCTTGGATTAACAAGGGGAAATGTTTACTGCACTTCTATC 2400
Db 2341 AGGTCAATTCATGGGAGTTGCTTGGATTAACAAGGGGAAATGTTTACTGCACTTCTATC 2400

Db 4561 TTTTGGCCCAACTGTATAGAAAGGATCGGTTACTTTTGGGTAGATACATACACA 4620
Qy 4621 CAGTATAGAGGAGCAATAGTGTATTATGATCTTTTGAAGTTGATCTCCAAACTTGGGCG 4680
Db 4621 CCAATGATAGAGGAGCAATAGTGTATTATGATCTTTTGAAGTTGATCTCCAAACTTGGGCG 4680
Qy 4681 AAAACATATCAGGATATGATCAGTATAGAGGTTAAGGTTGCTTTTGGCGATGTTGC 4740
Db 4681 AAAACATATCAGGATATGATCAGTATAGAGGTTAAGGTTGCTTTTGGCGATGTTGC 4740
Qy 4741 TTGTTCCGCTCGGAAACTGTTGCTTAGGCTTTCCGAGCTGAAACGAGCTATCAAGGAGT 4800
Db 4741 TTGTTCCGCTCGGAAACTGTTGCTTAGGCTTTCCGAGCTGAAACGAGCTATCAAGGAGT 4800
Qy 4801 TCAATAAAACCGGATGATGTTGCTTTGCTTTTAAATGTTTAAACAAATTTTGTGTGA 4860
Db 4801 TCAATAAAACCGGATGATGTTGCTTTGCTTTTAAATGTTTAAACAAATTTTGTGTGA 4860
Qy 4861 TAAATTTTATTTAGAACTTTGTTTAAAGGTTTAAAGGTTTAAAGGTTTAAAGGTTT 4920
Db 4861 TAAATTTTATTTAGAACTTTGTTTAAAGGTTTAAAGGTTTAAAGGTTTAAAGGTTT 4920
Qy 4921 TTAGCGAGTTCAATTGATCTTTTCCGAAACAGGATGAGATCTTCCGCAATTCATGACTAAGG 4980
Db 4921 TTAGCGAGTTCAATTGATCTTTTCCGAAACAGGATGAGATCTTCCGCAATTCATGACTAAGG 4980
Qy 4981 TCAAGAGTTTGAATATCGACTGTGGA CAGATTAATGCTGTTTAAAGATGATAGTCTTT 5040
Db 4981 TCAAGAGTTTGAATATCGACTGTGGA CAGATTAATGCTGTTTAAAGATGATAGTCTTT 5040
Qy 5041 CTGATGTAGATTTACTTAAAGGTTTAAAGGTTTAAAGGTTTAAAGGTTTAAAGGTTT 5100
Db 5041 CTGATGTAGATTTACTTAAAGGTTTAAAGGTTTAAAGGTTTAAAGGTTTAAAGGTTT 5100
Qy 5101 ATTTCGTAGTGTCTGGGAGTGGATCTCCCGGATAACTCCCGGATACTCCCGGTTT 5160
Db 5101 ATTTCGTAGTGTCTGGGAGTGGATCTCCCGGATAACTCCCGGATACTCCCGGTTT 5160
Qy 5161 GTATTTGATAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 5220
Db 5161 GTATTTGATAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAGTAAG 5220
Qy 5221 CTGCTTGCACAAAAGAAATTTTCTTTTAAAGCTAAATCCCTTAATTTTCAATACATCCGAGG 5280
Db 5221 CTGCTTGCACAAAAGAAATTTTCTTTTAAAGCTAAATCCCTTAATTTTCAATACATCCGAGG 5280
Qy 5281 ATGCTGAGAGCAACCCGTTGCAAGTGTATGATGATATCAAGAGAGTGGCTATGGAAGAAG 5340
Db 5281 ATGCTGAGAGCAACCCGTTGCAAGTGTATGATGATATCAAGAGAGTGGCTATGGAAGAAG 5340
Qy 5341 GATACGTCTTTTATCTTTGAGTTCGTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 5400
Db 5341 GATACGTCTTTTATCTTTGAGTTCGTTTCAATTTTCAATTTTCAATTTTCAATTTTCAAT 5400
Qy 5401 GAAAAGGTTTGAAGGAACTGTTTGAAGTGTGACAGACGCTCCGCAATTTGAATCACTG 5460
Db 5401 GAAAAGGTTTGAAGGAACTGTTTGAAGTGTGACAGACGCTCCGCAATTTGAATCACTG 5460
Qy 5461 AAAAGGTTTGAAGGAGTTCGTTGATCAAGTACCAATGGCTGTGAACCTGAAAAGGTTTC 5520
Db 5461 AAAAGGTTTGAAGGAGTTCGTTGATCAAGTACCAATGGCTGTGAACCTGAAAAGGTTTC 5520
Qy 5521 CGGAAAACAAAAGAAATGTTAGTAAATGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTT 5580
Db 5521 CGGAAAACAAAAGAAATGTTAGTAAATGTTTAAATGTTTAAATGTTTAAATGTTTAAATGTT 5580
Qy 5581 GTAAGAGGTTTAAATGTTGAGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGG 5640
Db 5581 GTAAGAGGTTTAAATGTTGAGAAATGAGGAAATGAGGAAATGAGGAAATGAGGAAATGAGG 5640
Qy 5641 CGTCATCGATGCTTTTAAATGATATGCTTTATCAATCAATCAATCAATCAATCAATCAATCAAT 5700

Db 5641 CGTCATCGATGCTTTTAAATCAATATGCTTTATACAACTCAATCAATCAATCAATCAAT 5700
Qy 5701 TTACTTATCTTTCCGCTTACGAGATCTCTGTCAGCTGATCAATCTGTGTACAATGCAATT 5760
Db 5701 TTACTTATCTTTCCGCTTACGAGATCTCTGTCAGCTGATCAATCTGTGTACAATGCAATT 5760
Qy 5761 GGGTAAACAGTTTCAAACGCAACAGCTAGAGCAACAGTCCAAACAGCAATTTGCGGATGC 5820
Db 5761 GGGTAAACAGTTTCAAACGCAACAGCTAGAGCAACAGTCCAAACAGCAATTTGCGGATGC 5820
Qy 5821 CTGGAACCTGTCGCTAGTATGACAGTGAAGATTTCTTGGCATCGATTTCTATGTGTATAG 5880
Db 5821 CTGGAACCTGTCGCTAGTATGACAGTGAAGATTTCTTGGCATCGATTTCTATGTGTATAG 5880
Qy 5881 ATATAATTCGACGCTGATCGTTGATCAGGCGTTTAAATAGCTTTGATATAGATA 5940
Db 5881 ATATAATTCGACGCTGATCGTTGATCAGGCGTTTAAATAGCTTTGATATAGATA 5940
Qy 5941 TAGAATATAGAGTTGATTAATCAACCCGCAACGCAATCTACTGAAATCGTTAAACGCGAC 6000
Db 5941 TAGAATATAGAGTTGATTAATCAACCCGCAACGCAATCTACTGAAATCGTTAAACGCGAC 6000
Qy 6001 TCAGAGGTTAGACGATCTACTGATGATTAAGGGCTTCAATCAATTAATTTGGCTAATGA 6060
Db 6001 TCAGAGGTTAGACGATCTACTGATGATTAAGGGCTTCAATCAATTAATTTGGCTAATGA 6060
Qy 6061 ACTGGTTTCGTGGAACTGGCATGTTCAATCAACGAGGCTTTGAGACTGCTAGTGGACTGT 6120
Db 6061 ACTGGTTTCGTGGAACTGGCATGTTCAATCAACGAGGCTTTGAGACTGCTAGTGGACTGT 6120
Qy 6121 CTGGAACCAACTCCGCTACTTAGCTATTGTTGTGAGATTTCTTAAATTAAGTCCGCTG 6180
Db 6121 CTGGAACCAACTCCGCTACTTAGCTATTGTTGTGAGATTTCTTAAATTAAGTCCGCTG 6180
Qy 6181 AAGACTTAAATTAAGTGGCTGATCAACCAATCAAGAGGCTTTGAGACTGCTAGTGGACTGT 6240
Db 6181 AAGACTTAAATTAAGTGGCTGATCAACCAATCAAGAGGCTTTGAGACTGCTAGTGGACTGT 6240
Qy 6241 TATAACGATTTGTCATATCTGGATCCCAACAGTTTAAACCATGTTGATGTTGTTGTTA 6300
Db 6241 TATAACGATTTGTCATATCTGGATCCCAACAGTTTAAACCATGTTGATGTTGTTGTTA 6300
Qy 6301 TGGCGTTAAACATCGAGAGGTTGGAATCTCCCTTAACCGCGGTAGCGGCCA 6355
Db 6301 TGGCGTTAAACATCGAGAGGTTGGAATCTCCCTTAACCGCGGTAGCGGCCA 6355

RESULT 2

AX795380 LOCUS 6355 bp mRNA linear PAT 04-OCT-2003
DEFINITION Sequence 7 from Patent WO03052108.
ACCESSION AX795380
VERSION AX795380.1 GI:37516053
KEYWORDS synthetic construct
SOURCE synthetic construct
ORGANISM synthetic construct
REFERENCE 1
AUTHORS Metzlauff, M. H., Gosselle, V. M., Meulewaeter, F. and Fache, J. C.
TITLE Improved methods and means for delivering inhibitory rna to plants
and applications thereof
JOURNAL Patent: WO 03052108-A 7 26-JUN-2003;
Bayer BioScience N.V. (BE)
FEATURES
source Location/Qualifiers
1. .6355
/organism="synthetic construct"
/mol_type="mRNA"
/db_xref="taxon:32630"
/note="cDNA sequence of the genome of TMV-U2"

ORIGIN

Query Match 100.0%; Score 6355; DB 6; Length 6355;
Best Local Similarity 100.0%; Pred. No. 0;

Matches 6355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;			
Qy	1	GATGTTTTTAATAGTTTTGCAACAACAATTAABAAACAAAACAAACATATTACAAAACA	60
Db	1	GATGTTTTTAATAGTTTTGCAACAACAATTAABAAACAAAACAAACATATTACAAAACA	60
Qy	61	AACAAACAACAAATGGCACACATACAATCTATAATTTAGCAAGCGCCCTCTTCAAGAACGGTGAG	120
Db	61	AACAAACAACAAATGGCACACATACAATCTATAATTTAGCAAGCGCCCTCTTCAAGAACGGTGAG	120
Qy	121	TGGTAAAAACACTCTCTGTTTAATGACCTTTGCAAGAGCGGCATGTAACGATACGGCCGTGGA	180
Db	121	TGGTAAAAACACTCTCTGTTTAATGACCTTTGCAAGAGCGGCATGTAACGATACGGCCGTGGA	180
Qy	181	AGAAATTTAAACCGCGACCGTAGACCAAGAGTCAACTTTTCCAAACACTATTACGAGAGA	240
Db	181	AGAAATTTAAACCGCGACCGTAGACCAAGAGTCAACTTTTCCAAACACTATTACGAGAGA	240
Qy	241	GCAAAACGCTTCTAGTCTTCCAAACCGTACCGGAGTTCAGATTTACCTTTTATAATACTCA	300
Db	241	GCAAAACGCTTCTAGTCTTCCAAACCGTACCGGAGTTCAGATTTACCTTTTATAATACTCA	300
Qy	301	AAATGCGGTACACAGTTTGGCTGGAGTTTGGAGAGCAATTAGAAATTTGGAATATCTGATGCT	360
Db	301	AAATGCGGTACACAGTTTGGCTGGAGTTTGGAGAGCAATTAGAAATTTGGAATATCTGATGCT	360
Qy	361	ACAAGTTCCCTATGGATCGCGACATATGATATAGTGGGAACCTTTGCGACACATTTGTT	420
Db	361	ACAAGTTCCCTATGGATCGCGACATATGATATAGTGGGAACCTTTGCGACACATTTGTT	420
Qy	421	CAAAAGCAGGATTTACGTGCATTTCTGATGCCCAATCTGCAATACAGATATATATGAG	480
Db	421	CAAAAGCAGGATTTACGTGCATTTCTGATGCCCAATCTGCAATACAGATATATATGAG	480
Qy	481	GCAAGAGGACAAAGAGACTCAATTTGAGATGTATTTGTCAGATTTGTCCTCTAAACA	540
Db	481	GCAAGAGGACAAAGAGACTCAATTTGAGATGTATTTGTCAGATTTGTCCTCTAAACA	540
Qy	541	GGTAATTCCTGAGTTTCAAGGGAGGCTTTTAAACAGGTATGCAAGCTCCCAACGAGT	600
Db	541	GGTAATTCCTGAGTTTCAAGGGAGGCTTTTAAACAGGTATGCAAGCTCCCAACGAGT	600
Qy	601	CTGCTGCTCTAAACTTTTCAAGGATCTCGAATACATCCCGCAGAGAAATGCTGAGAAG	660
Db	601	CTGCTGCTCTAAACTTTTCAAGGATCTCGAATACATCCCGCAGAGAAATGCTGAGAAG	660
Qy	661	ATAGCGTTGCTCTGCACAGTTTGTATGATATTTCTGTGTCATGAGTTTGGAGCTGCGTT	720
Db	661	ATAGCGTTGCTCTGCACAGTTTGTATGATATTTCTGTGTCATGAGTTTGGAGCTGCGTT	720
Qy	721	AATATCTAAGAAATATACATGATATGATATGATATTTTCCGTGTCATGAGTTTGGCAGAGTATTA	780
Db	721	AATATCTAAGAAATATACATGATATGATATGATATTTTCCGTGTCATGAGTTTGGCAGAGTATTA	780
Qy	781	AGACCCAGCGGTTACGTTTAAATGAAATAGGCGCAACTTTTCAAAAGAGAGGTGATGA	840
Db	781	AGACCCAGCGGTTACGTTTAAATGAAATAGGCGCAACTTTTCAAAAGAGAGGTGATGA	840
Qy	841	TGTTTCTTTTCTTCTGATGAAAGTACTTTTAAATTTAGTATATAAATACAAAATAT	900
Db	841	TGTTTCTTTTCTTCTGATGAAAGTACTTTTAAATTTAGTATATAAATACAAAATAT	900
Qy	901	CTTGCAATATGATGTTAAATCTTACTTTCTGCTTCTAGTAGAATAGTTTACTTTAAGGA	960
Db	901	CTTGCAATATGATGTTAAATCTTACTTTCTGCTTCTAGTAGAATAGTTTACTTTAAGGA	960
Qy	961	ATTTTGTAGTCACTAGGTTTAACTCTGTTTGTAAATTTTACCAAGTAGATACCTATAT	1020
Db	961	ATTTTGTAGTCACTAGGTTTAACTCTGTTTGTAAATTTTACCAAGTAGATACCTATAT	1020
Qy	1021	TCTGTACAAGAGTCTTAGACAAGTAGGTTGATGATGATCAGTTCTATGAGGCGATGGA	1080
Db	1021	TCTGTACAAGAGTCTTAGACAAGTAGGTTGATGATGATCAGTTCTATGAGGCGATGGA	1080

Qy	1081	AGACGCTTTGCTTACAGAAAACCTTGGCCATGTTTCAACACATGAAAGAGCAATCTTTAG	1140
Db	1081	AGACGCTTTGCTTACAGAAAACCTTGGCCATGTTTCAACACATGAAAGAGCAATCTTTAG	1140
Qy	1141	AGACACGGCTTCGGTTAACTTTTGGTTCCCTAAAGATGAAGGACATGGTGAATGACGCT	1200
Db	1141	AGACACGGCTTCGGTTAACTTTTGGTTCCCTAAAGATGAAGGACATGGTGAATGACGCT	1200
Qy	1201	GTTTGAAGGTTCTATTATACAGCAAAAAGATGACAGAGGTGAGTCAATTTGTTAATCGTGA	1260
Db	1201	GTTTGAAGGTTCTATTATACAGCAAAAAGATGACAGAGGTGAGTCAATTTGTTAATCGTGA	1260
Qy	1261	CTTTCGTTTACACAGTCTTAATCATATCAGAACATATCAAGCCAAAGCGTTAACTTACCA	1320
Db	1261	CTTTCGTTTACACAGTCTTAATCATATCAGAACATATCAAGCCAAAGCGTTAACTTACCA	1320
Qy	1321	GAAACGTTATCTTTCTGCGAGTCTATTAAGATCCCGCGTGATTAATCAATGGTGTACTGC	1380
Db	1321	GAAACGTTATCTTTCTGCGAGTCTATTAAGATCCCGCGTGATTAATCAATGGTGTACTGC	1380
Qy	1381	TAGGTCGTAATGGGATGATATAAGCAATTTCTTCAACCTTGTCAATGACTTTCTTCTT	1440
Db	1381	TAGGTCGTAATGGGATGATATAAGCAATTTCTTCAACCTTGTCAATGACTTTCTTCTT	1440
Qy	1441	GCAGACTTAAGCTGCGCTTCAAGACGATATAGTAATGGGAAAGTTTGGTCTTGGGA	1500
Db	1441	GCAGACTTAAGCTGCGCTTCAAGACGATATAGTAATGGGAAAGTTTGGTCTTGGGA	1500
Qy	1501	TAAGACCACTTCTGAACTTATTTGGGATGAGGTGGGCAAAATTTTGGGAAACGTTTTCCC	1560
Db	1501	TAAGACCACTTCTGAACTTATTTGGGATGAGGTGGGCAAAATTTTGGGAAACGTTTTCCC	1560
Qy	1561	CACATCAAGAGAGATTTGGTGAGAGGAAATTTCTGGATGTAAAGTGAAGATGCTCTGAA	1620
Db	1561	CACATCAAGAGAGATTTGGTGAGAGGAAATTTCTGGATGTAAAGTGAAGATGCTCTGAA	1620
Qy	1621	GATCAAGATCCAGATCTGATGTGCATGCAAGAGACAGGTTCTGAGTCAATACACAA	1680
Db	1621	GATCAAGATCCAGATCTGATGTGCATGCAAGAGACAGGTTCTGAGTCAATACACAA	1680
Qy	1681	GTCTGAGGAGTTACCGCATCTAGATATCAAGAAAGGACTTAGAAGAAAGCTGAGCAAAATGTA	1740
Db	1681	GTCTGAGGAGTTACCGCATCTAGATATCAAGAAAGGACTTAGAAGAAAGCTGAGCAAAATGTA	1740
Qy	1741	CGACGCTTATCAGAAATATCTATCTTAAAGGTGCTGATATAATTCGATATCGGAAAGTT	1800
Db	1741	CGACGCTTATCAGAAATATCTATCTTAAAGGTGCTGATATAATTCGATATCGGAAAGTT	1800
Qy	1801	CAAGACATGTGCAAGGCTTTAGATGTTAGTCTGATGTGGCAGCAGAGTAATCGTTGC	1860
Db	1801	CAAGACATGTGCAAGGCTTTAGATGTTAGTCTGATGTGGCAGCAGAGTAATCGTTGC	1860
Qy	1861	AGTGGCGGAGAAATAGAAAGCGGTTTAACTCTTACTTTTGAAGCAACCGAGGAGAAATGT	1920
Db	1861	AGTGGCGGAGAAATAGAAAGCGGTTTAACTCTTACTTTTGAAGCAACCGAGGAGAAATGT	1920
Qy	1921	GGCTAAGGCTCTTAAAGCAACGCGCTCTGAGGCGGTGATGTCTTGAACCGACATCCGA	1980
Db	1921	GGCTAAGGCTCTTAAAGCAACGCGCTCTGAGGCGGTGATGTCTTGAACCGACATCCGA	1980
Qy	1981	AGAGGTGAACGTAATAATAATTTCTATTGCTGAGAAAGGAGATTTGCTGTGTGTCAGA	2040
Db	1981	AGAGGTGAACGTAATAATAATTTCTATTGCTGAGAAAGGAGATTTGCTGTGTGTCAGA	2040
Qy	2041	AAGTCATGTTTTCACGAATGCTAACTTATAGACACCGAGAGTTGGAGTCCCTCAACGATTT	2100
Db	2041	AAGTCATGTTTTCACGAATGCTAACTTATAGACACCGAGAGTTGGAGTCCCTCAACGATTT	2100
Qy	2101	CCATAAGGCTTGGTGATAGTGTGATATCAAAAGCAAAATGGCATCGGTTGCTACACTGG	2160
Db	2101	CCATAAGGCTTGGTGATAGTGTGATATCAAAAGCAAAATGGCATCGGTTGCTACACTGG	2160

QY 2161 CTCACCTCAAGTTTCAACAAATGAAGAACTATGTGCACAGTTTGGCAGCTTCGTGTCCGC 2220
Db 2161 CTCACCTCAAGTTTCAACAAATGAAGAACTATGTGCACAGTTTGGCAGCTTCGTGTCCGC 2220
QY 2221 CACTGTATCAAAATCTATCAAGTCACTAAAGGATGAAGTCGGGTATGATTCGTATCCAG 2280
Db 2221 CACTGTATCAAAATCTATCAAGTCACTAAAGGATGAAGTCGGGTATGATTCGTATCCAG 2280
QY 2281 GGAGAAAGTTGGTGTGGGATGTCACTTTGAAAGAGTGGCTCTCAAACTCGGGCCAA 2340
Db 2281 GGAGAAAGTTGGTGTGGGATGTCACTTTGAAAGAGTGGCTCTCAAACTCGGGCCAA 2340
QY 2341 AGGTCAATTCATGGGGAGTTGTCCTGGATTACAAGGGGAAATGTTTACTGACATTCATC 2400
Db 2341 AGGTCAATTCATGGGGAGTTGTCCTGGATTACAAGGGGAAATGTTTACTGACATTCATC 2400
QY 2401 TTATGAAGGAGATAGAAATGGTGACTGAGAGCGACTGGAGGAGGTGGCTGTATCATCTGA 2460
Db 2401 TTATGAAGGAGATAGAAATGGTGACTGAGAGCGACTGGAGGAGGTGGCTGTATCATCTGA 2460
QY 2461 TACAATGGTATATCTGATATTCGAAAGCTCCAAAATCTGAGGAAACAAATGAGAGCGG 2520
Db 2461 TACAATGGTATATCTGATATTCGAAAGCTCCAAAATCTGAGGAAACAAATGAGAGCGG 2520
QY 2521 TGAACCCCAACGACTACTGCAAGATGGTACTTGTGGATGGGTGGCTGGTGTGAAA 2580
Db 2521 TGAACCCCAACGACTACTGCAAGATGGTACTTGTGGATGGGTGGCTGGTGTGAAA 2580
QY 2581 GTACAAAGGAGATTTGAAAGATTTGATCTTGATGAGGATTTGATCTTGGTTCCTGGAAA 2640
Db 2581 GTACAAAGGAGATTTGAAAGATTTGATCTTGATGAGGATTTGATCTTGGTTCCTGGAAA 2640
QY 2641 ACAAGCTGCTGATGATCAGAAGAGGCTTAATTCATCTGGAATGAAGAGCCAAAT 2700
Db 2641 ACAAGCTGCTGATGATCAGAAGAGGCTTAATTCATCTGGAATGAAGAGCCAAAT 2700
QY 2701 GGACAAATGTGAGAACGGTAGATTCACTTCTAATGCATCCAAAACCGGATCACAAGAG 2760
Db 2701 GGACAAATGTGAGAACGGTAGATTCACTTCTAATGCATCCAAAACCGGATCACAAGAG 2760
QY 2761 GCTTTTATGATCAAGGTTGATGCTGCACACCGTTGTGTTAACTTCTCTGGTGTAT 2820
Db 2761 GCTTTTATGATCAAGGTTGATGCTGCACACCGTTGTGTTAACTTCTCTGGTGTAT 2820
QY 2821 CTCTGGTGGACATCGCATACATTTACGGAGATACACAGCAGATTCCTTTCAATTAACAG 2880
Db 2821 CTCTGGTGGACATCGCATACATTTACGGAGATACACAGCAGATTCCTTTCAATTAACAG 2880
QY 2881 AGTTCAGAAATTCCTCGTATCCCAACATTTTGAGAACTGCAAGTGGATGAAGTTGAGAT 2940
Db 2881 AGTTCAGAAATTCCTCGTATCCCAACATTTTGAGAACTGCAAGTGGATGAAGTTGAGAT 2940
QY 2941 GAGGAGGACACACTGAGATGCCAGGTGATGTGAAATTTTCTCTACAATCGAAGTACGA 3000
Db 2941 GAGGAGGACACACTGAGATGCCAGGTGATGTGAAATTTTCTCTACAATCGAAGTACGA 3000
QY 3001 AGGACGGGTGACACCACTTCAACTGTACACAGATCGGTCTCACTCGAGATGATAGCGG 3060
Db 3001 AGGACGGGTGACACCACTTCAACTGTACACAGATCGGTCTCACTCGAGATGATAGCGG 3060
QY 3061 TAAGGGAGTACTAAACAGTGTTCCTCAACCACTTAAGAGGAAATTTGTAATCTTCACTCA 3120
Db 3061 TAAGGGAGTACTAAACAGTGTTCCTCAACCACTTAAGAGGAAATTTGTAATCTTCACTCA 3120
QY 3121 GGCTGATAAATTTGAGTTAGAGGAGAGGCTATAAGAAATGTGAACCGTTCAATGAGAT 3180
Db 3121 GGCTGATAAATTTGAGTTAGAGGAGAGGCTATAAGAAATGTGAACCGTTCAATGAGAT 3180
QY 3181 CCAAGGAGAAACCTTTGAAGATGTGCTGGTCAAGATTTGACGGCAACTCCACTGACTCT 3240
Db 3181 CCAAGGAGAAACCTTTGAAGATGTGCTGGTCAAGATTTGACGGCAACTCCACTGACTCT 3240
QY 3241 GATTTCCAGTCTTCCCGCATGTTTCTAGTCGCTCTGACTAGACACAAAAGAGCTTCAA 3300

Db 3341 GATTTCCAGTCTTCCCGCATGTTTCTAGTCGCTCTGACTAGACACAAAAGAGCTTCAA 3300
QY 3301 ATATTACACCGTAGTGTAGATCCTTTAGTACAGATAAATAGTGAATTTGCTTCTTTAAG 3360
Db 3301 ATATTACACCGTAGTGTAGATCCTTTAGTACAGATAAATAGTGAATTTGCTTCTTTAAG 3360
QY 3361 CTCCTTCTTTTGAAGATATATGTGTAAGCAGGTAGTAGATAGCAATTTACAGATGGA 3420
Db 3361 CTCCTTCTTTTGAAGATATATGTGTAAGCAGGTAGTAGATAGCAATTTACAGATGGA 3420
QY 3421 TGCAGTGTTCARAGGTCAATAATCTCTTTTGGGCAACCTTAATCAGAGACTTTCCAGA 3480
Db 3421 TGCAGTGTTCARAGGTCAATAATCTCTTTTGGGCAACCTTAATCAGAGACTTTCCAGA 3480
QY 3481 TCTACAGTCTTATTTACAGATGTATGCCCTCCCTGGTAAATAGTACTATCTTAAACAAGTATGA 3540
Db 3481 TCTACAGTCTTATTTACAGATGTATGCCCTCCCTGGTAAATAGTACTATCTTAAACAAGTATGA 3540
QY 3541 TGCTGTTTACCATGAGGTTACGTGATATATAGTCTTAAATGTGAAGATTTGTCTTTGATTT 3600
Db 3541 TGCTGTTTACCATGAGGTTACGTGATATATAGTCTTAAATGTGAAGATTTGTCTTTGATTT 3600
QY 3601 TTTCCAAAAGTATTTCCGATGTCGAAAGGAGGTGAACCAATGCTAGAGCCAGTTTTCGCTAC 3660
Db 3601 TTTCCAAAAGTATTTCCGATGTCGAAAGGAGGTGAACCAATGCTAGAGCCAGTTTTCGCTAC 3660
QY 3661 CGGGGCGGAACCGCAAGGGCTGCAGGACTACTCGAAAAATCTGGTTGCAATGATTAAGAAG 3720
Db 3661 CGGGGCGGAACCGCAAGGGCTGCAGGACTACTCGAAAAATCTGGTTGCAATGATTAAGAAG 3720
QY 3721 AAATTTCAACGCAACGACCTGACGGGACGATGTGACATTTGAGAGCAACCGCATCTGTGT 3780
Db 3721 AAATTTCAACGCAACGACCTGACGGGACGATGTGACATTTGAGAGCAACCGCATCTGTGT 3780
QY 3781 AGTAGATAAGTATTTTGTAGTACTATTTTATTAAGAAAGAAATACACAAAATATATTCG 3840
Db 3781 AGTAGATAAGTATTTTGTAGTACTATTTTATTAAGAAAGAAATACACAAAATATATTCG 3840
QY 3841 TGCAAGTATGACGAAGGATTTCAATGATGAGATGGTTGGAAGAAAGAGTACTATTT 3900
Db 3841 TGCAAGTATGACGAAGGATTTCAATGATGAGATGGTTGGAAGAAAGAGTACTATTT 3900
QY 3901 GGAAGCTTGGCTTAACTACAAATTTTACAGATCTGCCGGCCATCGATCAGTCAAGCAAT 3960
Db 3901 GGAAGCTTGGCTTAACTACAAATTTTACAGATCTGCCGGCCATCGATCAGTCAAGCAAT 3960
QY 3961 GATCAAGGCTCAACCAAAACAGAAATTTGACCTTTCAATTCAGAAATGAAATACCTGCTCT 4020
Db 3961 GATCAAGGCTCAACCAAAACAGAAATTTGACCTTTCAATTCAGAAATGAAATACCTGCTCT 4020
QY 4021 GCAAAACAATTTGCTACCATTTGGAAGCAGATCAACGGTATTTTGGCGGTTTCTCAGAGCT 4080
Db 4021 GCAAAACAATTTGCTACCATTTGGAAGCAGATCAACGGTATTTTGGCGGTTTCTCAGAGCT 4080
QY 4081 TACAAGTTGCTGCTCGAGGCAATTTGATTTCTAAGAAAGTTTCTTTTCTTTACTAGGAAAC 4140
Db 4081 TACAAGTTGCTGCTCGAGGCAATTTGATTTCTAAGAAAGTTTCTTTTCTTTACTAGGAAAC 4140
QY 4141 TCCAGAACAGATTCAGAAATTTTCTCGGATCTCGACTCGCACTGCTTCTTATGATGTGTT 4200
Db 4141 TCCAGAACAGATTCAGAAATTTTCTCGGATCTCGACTCGCACTGCTTCTTATGATGTGTT 4200
QY 4201 AGAACTGGATATTTCTAAGATGATGATAAGTACAGAAACGAGTTTCAATGTCTGTAGAGTA 4260
Db 4201 AGAACTGGATATTTCTAAGATGATGATAAGTACAGAAACGAGTTTCAATGTCTGTAGAGTA 4260
QY 4261 TGAATATGGAAGATTTGGGTCTCAATGATTTTGGCCGAAGTGTGGAACAAGGGCA 4320
Db 4261 TGAATATGGAAGATTTGGGTCTCAATGATTTTGGCCGAAGTGTGGAACAAGGGCA 4320
QY 4321 CAGGAAAACACTTTGAAGGATTTACATTTGCTGGAATTCAGACATGCTGTGGTATCAAG 4380

Db	4321	CAGGAACAACCTTTGAAGGATTAACAATTGCTGGGAATCAAGACATGTCGTGTGATTAACAAG	4380
Qy	4381	GAAGAAGCGGTGATGTGACTACTTTTCATCGGCAATACGTGTTATAATACGACGTGCTTGGG	4440
Db	4381	GAAGAAGCGGTGATGTGACTACTTTTCATCGGCAATACGTGTTATAATACGACGTGCTTGGG	4440
Qy	4441	TTCAATGTTACCGATGGAAAAGGTCATAAAGAGTGCTTTTGTGGAGACGATTCCTCGTTTT	4500
Db	4441	TTCAATGTTACCGATGGAAAAGGTCATAAAGAGTGCTTTTGTGGAGACGATTCCTCGTTTT	4500
Qy	4501	GTAATTTTCCAAAGGTTTGGATTTTCCCTGCACATTCAGTCAATGCTAAATCTCAATGGAA	4560
Db	4501	GTAATTTTCCAAAGGTTTGGATTTTCCCTGCACATTCAGTCAATGCTAAATCTCAATGGAA	4560
Qy	4561	TTTTGAGGCCAAAATCTGATAGAAAGAGGTACGGTTACTTTTGTGGTAGATACATCATACA	4620
Db	4561	TTTTGAGGCCAAAATCTGATAGAAAGAGGTACGGTTACTTTTGTGGTAGATACATCATACA	4620
Qy	4621	CCATGATAAGGAGCAATAGTGTAATATGATCCTTTGAGTTGATCTCCAAAATCTTGGGGC	4680
Db	4621	CCATGATAAGGAGCAATAGTGTAATATGATCCTTTGAGTTGATCTCCAAAATCTTGGGGC	4680
Qy	4681	AAAACATATCAAGCATATGATCACTAGTAGAGAGTTAAGGTTGCTTTGTGCGATGTTGC	4740
Db	4681	AAAACATATCAAGCATATGATCACTAGTAGAGAGTTAAGGTTGCTTTGTGCGATGTTGC	4740
Qy	4741	TTGTTTCGCTCGAAAATCGTGCTTAGGCTTTCCCGACGTGAACGACGCTATCAAGGAGGT	4800
Db	4741	TTGTTTCGCTCGAAAATCGTGCTTAGGCTTTCCCGACGTGAACGACGCTATCAAGGAGGT	4800
Qy	4801	TCATAAAAACCGCGATGATGGTTCGTTTGTCTTTTAAATGTGTAAACAAAATTTTGTGTGA	4860
Db	4801	TCATAAAAACCGCGATGATGGTTCGTTTGTCTTTTAAATGTGTAAACAAAATTTTGTGTGA	4860
Qy	4861	TAAATTTTATTAGAACTTTGTTTTTAAATGGCTGTGTAGTCTCAGAGATACGTCAAAA	4920
Db	4861	TAAATTTTATTAGAACTTTGTTTTTAAATGGCTGTGTAGTCTCAGAGATACGTCAAAA	4920
Qy	4921	TTACGAGTTCATTTGATCTTTTCGAAAACAGGATGAGATACTTTCCGGCAATTCATGACTAAGG	4980
Db	4921	TTACGAGTTCATTTGATCTTTTCGAAAACAGGATGAGATACTTTCCGGCAATTCATGACTAAGG	4980
Qy	4981	TCAAGAGTGTAGAAATACGACTGTGCAACAGATTTATGGCTGTTTAAAGATGATGCTTTT	5040
Db	4981	TCAAGAGTGTAGAAATACGACTGTGCAACAGATTTATGGCTGTTTAAAGATGATGCTTTT	5040
Qy	5041	CTGATCTAGATTACTTTAAAGGTGTTAACTTAGTTAGTAAAGAGGTATGTGTCTAGCTG	5100
Db	5041	CTGATCTAGATTACTTTAAAGGTGTTAACTTAGTTAGTAAAGAGGTATGTGTCTAGCTG	5100
Qy	5101	ATTTGGTAGTGTCTGGGAGTGAATCTCCCGATAACTCGCCGTGGTGGTGCAGTGTGTT	5160
Db	5101	ATTTGGTAGTGTCTGGGAGTGAATCTCCCGATAACTCGCCGTGGTGGTGCAGTGTGTT	5160
Qy	5161	GTAATTTAGATAAGAGAAATGAAAAGGAGTAAAGGAAGCAACGCTGGGTGCGTATCACGCC	5220
Db	5161	GTAATTTAGATAAGAGAAATGAAAAGGAGTAAAGGAAGCAACGCTGGGTGCGTATCACGCC	5220
Qy	5221	CTGCTTGCAAAAGAAATTTTTTCTTTTAAAGCTAATCCCTAAATTAATCAATAACATCCGAGG	5280
Db	5221	CTGCTTGCAAAAGAAATTTTTTCTTTTAAAGCTAATCCCTAAATTAATCAATAACATCCGAGG	5280
Qy	5281	ATGCTGAGAGCACCCGTGGCAAGTGTAGTGAATATCAAGAGGATGGCTATCGAGAGAG	5340
Db	5281	ATGCTGAGAGCACCCGTGGCAAGTGTAGTGAATATCAAGAGGATGGCTATCGAGAGAG	5340
Qy	5341	GATACTGTCTTTATCTTTTGAGTTCGTTTCAAATTTGTGTAGTACATAAAAAATAATGTAA	5400
Db	5341	GATACTGTCTTTATCTTTTGAGTTCGTTTCAAATTTGTGTAGTACATAAAAAATAATGTAA	5400
Qy	5401	GAAGAAGTTGAGGAAACGTAATTTTGTAGTGTGACAGACGGCTCGCCAAATTTGAATCACTG	5460
Db	5401	GAAGAAGTTTGTAGGAAACGTAATTTTGTAGTGTGACAGACGGCTCGCCAAATTTGAATCACTG	5460

Qy	5461	AAAAAGGTGTTGAGGAGTTTCGTGGATGAAGTACCAATGGCTGTGAAACTCGAAAAAGGTTTC	5520
Db	5461	AAAAAGGTGTTGAGGAGTTTCGTGGATGAGTACCAATGGCTGTGAAACTCGAAAAAGGTTTC	5520
Qy	5521	CGGAAAACAAAAAGAAATGGTATAGTTAATATGTTAATAATAGAAAATAAATAACAGTG	5580
Db	5521	CGGAAAACAAAAAGAAATGGTATAGTTAATATGTTAATAATAGAAAATAAATAACAGAGT	5580
Qy	5581	GTAAAGAGGGTTTTAAAAATTGAGCAAAATTCAGGATATATGTAAGTGTATGACGAGTCTATCG	5640
Db	5581	GTAAAGAGGGTTTTAAAAATTGAGCAAAATTCAGGATATATGTAAGTGTATGACGAGTCTATCG	5640
Qy	5641	CGTCATCGAGTACGTTTTTAATCAATATGCTTTATACAATCAAATCTCTCCGAGCCAAATTTGT	5700
Db	5641	CGTCATCGAGTACGTTTTTAATCAATATGCTTTATACAATCAAATCTCTCCGAGCCAAATTTGT	5700
Qy	5701	TTACTTATCTTCGCTTACGCAATCCTGTGACGATGATCAATCTGTGTATCAAAATGCATTT	5760
Db	5701	TTACTTATCTTCGCTTACGCAATCCTGTGACGATGATCAATCTGTGTATCAAAATGCATTT	5760
Qy	5761	GGGTAAACAGTGTTCAAACGCAACAGCTAGGACAAAGTCCAAACAGCAATTTTCGGGATGC	5820
Db	5761	GGGTAAACAGTGTTCAAACGCAACAGCTAGGACAAAGTCCAAACAGCAATTTTCGGGATGC	5820
Qy	5821	CTGGAAACCTGTGCCTAGTATGACAGTGAATTTCTTGCAATCGGATTTCTATGTGTATAG	5880
Db	5821	CTGGAAACCTGTGCCTAGTATGACAGTGAATTTCTTGCAATCGGATTTCTATGTGTATAG	5880
Qy	5881	ATATAATTCAGCGCTTGATCCGTTTGATCAGCGGTTTATTAATAGCTTTTGATACTAGAAA	5940
Db	5881	ATATAATTCAGCGCTTGATCCGTTTGATCAGCGGTTTATTAATAGCTTTTGATACTAGAAA	5940
Qy	5941	TAGAAATAATAGAGTTTGATAATCAACCGCACCGAATACCTACTGAAATCGTTAACCGCAC	6000
Db	5941	TAGAAATAATAGAGTTTGATAATCAACCGCACCGAATACCTACTGAAATCGTTAACCGCAC	6000
Qy	6001	TCAGAGGTAGACGATGCTACTCTAGCTATAAGGGGCTTCAATCAATAATTTGGCTAATGA	6060
Db	6001	TCAGAGGTAGACGATGCTACTCTAGCTATAAGGGGCTTCAATCAATAATTTGGCTAATGA	6060
Qy	6061	ACTGGTTCTGGGAACTGGGCATGTTTCAATCAAGCAGGCTTTGAGACTGCTAGTGGACTTGT	6120
Db	6061	ACTGGTTCTGGGAACTGGGCATGTTTCAATCAAGCAGGCTTTGAGACTGCTAGTGGACTTGT	6120
Qy	6121	CTGGACCAACATCCGCTACTTAGCTATTGTTGTGAGATTTCTCTAAATAAAGATCGCTG	6180
Db	6121	CTGGACCAACATCCGCTACTTAGCTATTGTTGTGAGATTTCTCTAAATAAAGATCGCTG	6180
Qy	6181	AAGACTTTAAATTTCAAGGTGGCTGATACCAAAATCAGCAGTGGTGTGTTCTGCCACTTAAA	6240
Db	6181	AAGACTTTAAATTTCAAGGTGGCTGATACCAAAATCAGCAGTGGTGTGTTCTGCCACTTAAA	6240
Qy	6241	TATAACGATTTGTCATATCTGGATCCAAACAGTTAAACCATGTGATGTTGATCTGTGGTA	6300
Db	6241	TATAACGATTTGTCATATCTGGATCCAAACAGTTAAACCATGTGATGTTGATCTGTGGTA	6300
Qy	6301	TGGCGTAAACATCGAGAGTTTCGAATCCTCCCTAAACCGCCGGTAGCGGCCCA	6355
Db	6301	TGGCGTAAACATCGAGAGTTTCGAATCCTCCCTAAACCGCCGGTAGCGGCCCA	6355
RESULT 3			
TMGCC			
LOCUS	6355 bp ss-RNA linear VRL 03-AUG-1993		
DEFINITION	Tobacco mild green mosaic virus complete genome.		
ACCESSION	M34077.22483		
VERSION	M34077.1 GI:335243		
KEYWORDS	coat protein.		
SOURCE	Tobacco mild green mosaic virus		
ORGANISM	Tobacco mild green mosaic virus		
REFERENCE	Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.		
	1 (bases 6127 to 6355)		

Qy	481	GCACGAGGCAAAAGGACCTCAATTTGAGATGATTTTGTCCAGATTGTCCTCGTTCTCAACAA	540
Db	481	GCACGAGGCAAAAGGACCTCAATTTGAGATGATTTTGTCCAGATTGTCCTCGTTCTCAACAA	540
Qy	541	GGTAATTCCTGAGTTTCAAAAGGAGGCTTTTAAACAGGTATGCAGAACTCCCAACGAAAT	600
Db	541	GGTAATTCCTGAGTTTCAAAAGGAGGCTTTTAAACAGGTATGCAGAACTCCCAACGAAAT	600
Qy	601	CTGCTGCTCTAAACATTTTTCAGGATTCGAAATCATCCGCCAGAGAAATAGTGGTAGAAG	660
Db	601	CTGCTGCTCTAAACATTTTTCAGGATTCGAAATCATCCGCCAGAGAAATAGTGGTAGAAG	660
Qy	661	ATAGCGTTGCTCTGCACAGTTTGTATGATATATCTCTGTGCATGAGTTTGGAGCTGGTT	720
Db	661	ATAGCGTTGCTCTGCACAGTTTGTATGATATATCTCTGTGCATGAGTTTGGAGCTGGTT	720
Qy	721	AATATCTAAGAAATATACATGATGTTTGTATGATATATCTCTGTGCATGAGTTTGGAGCTGGTT	780
Db	721	AATATCTAAGAAATATACATGATGTTTGTATGATATATCTCTGTGCATGAGTTTGGAGCTGGTT	780
Qy	781	AGACGAGCGAGGTTTACGCTTAATGAATAGGCGCAACTTTTCAAAAGAGAGGTGATGA	840
Db	781	AGACGAGCGAGGTTTACGCTTAATGAATAGGCGCAACTTTTCAAAAGAGAGGTGATGA	840
Qy	841	TGTTCTCTTTTCTTCTGCTGATGAAGTACTTTTAAATATATAGTCAATAAATACAAAATAT	900
Db	841	TGTTCTCTTTTCTTCTGCTGATGAAGTACTTTTAAATATATAGTCAATAAATACAAAATAT	900
Qy	901	CTTGCAATATGATTAATCTTACTTTTCTGCTCTAGTAGAATAGTTTACITTAAGGA	960
Db	901	CTTGCAATATGATTAATCTTACTTTTCTGCTCTAGTAGAATAGTTTACITTAAGGA	960
Qy	961	ATTTTATAGTCACTAGGTTTAAATCTTGGTTTGTAAATTTTACCAAGTAGATACCTATAT	1020
Db	961	ATTTTATAGTCACTAGGTTTAAATCTTGGTTTGTAAATTTTACCAAGTAGATACCTATAT	1020
Qy	1021	TCTGTAACAAGTGTATAGCAAGTAGGTTGATGATGATCAGTTCTATGAGGCGATGGA	1080
Db	1021	TCTGTAACAAGTGTATAGCAAGTAGGTTGATGATGATCAGTTCTATGAGGCGATGGA	1080
Qy	1081	AGACGCTTTGCTTACAGAAACCTTGGCCATCTTCAACACTCAAGAGCAATCTTTTAC	1140
Db	1081	AGACGCTTTGCTTACAGAAACCTTGGCCATCTTCAACACTCAAGAGCAATCTTTTAC	1140
Qy	1141	AGACGCTTTGCTTACAGAAACCTTGGCCATCTTCAACACTCAAGAGCAATCTTTTAC	1200
Db	1141	AGACGCTTTGCTTACAGAAACCTTGGCCATCTTCAACACTCAAGAGCAATCTTTTAC	1200
Qy	1201	GTTTGAGGGTTCTATTACCAAGCAAAAGATGACAAAGGAGTGAATGTTTAAATCGTGA	1260
Db	1201	GTTTGAGGGTTCTATTACCAAGCAAAAGATGACAAAGGAGTGAATGTTTAAATCGTGA	1260
Qy	1261	CTTCGTTTACACAGTGTCTTATCATATCAGAAACATATCAAGCCAAAGCGTTTAACTTACCA	1320
Db	1261	CTTCGTTTACACAGTGTCTTATCATATCAGAAACATATCAAGCCAAAGCGTTTAACTTACCA	1320
Qy	1321	GAAAGTATATCTTTTCTGAGGAGTCTATAAGATCCGCGTGATATCAATGGTGTACTGC	1380
Db	1321	GAAAGTATATCTTTTCTGAGGAGTCTATAAGATCCGCGTGATATCAATGGTGTACTGC	1380
Qy	1381	TAGTCTGGAATGGATGATGAAGCAATTTCTTCAACCCCTTGTCAATGACTTTCTTCTT	1440
Db	1381	TAGTCTGGAATGGATGATGAAGCAATTTCTTCAACCCCTTGTCAATGACTTTCTTCTT	1440
Qy	1441	GCAGACTAAGCTGGCTGGCTTCAAGACGATATAGTAATGGGAAAGTTTCGGTCTGGA	1500
Db	1441	GCAGACTAAGCTGGCTGGCTTCAAGACGATATAGTAATGGGAAAGTTTCGGTCTGGA	1500
Qy	1501	TAAGACCACTTCTGAACTTATTTGGGATGAGTGGGCAAAATTTTGTGAAACGTTTCCC	1560
Db	1501	TAAGACCACTTCTGAACTTATTTGGGATGAGTGGGCAAAATTTTGTGAAACGTTTCCC	1560

Qy	1561	CACTATCAAGAGAGAGATTGGTGAGCAGGAAAAATTTCTGGATGTAAAGTGAGAAATCTCTGAA	1620
Db	1561	CACTATCAAGAGAGAGATTGGTGAGCAGGAAAAATTTCTGGATGTAAAGTGAGAAATCTCTGAA	1620
Qy	1621	GATCAAGATCCCAAGATCTGTATGTCACTGGAAGAAGACAGGTTCTGTAGCTGAATACACAA	1680
Db	1621	GATCAAGATCCCAAGATCTGTATGTCACTGGAAGAAGACAGGTTCTGTAGCTGAATACACAA	1680
Qy	1681	GTCTGAGGAGTTACCCATCTAGATATCAAGAGGACTTAAAGAAAGCTGAGCAAAATGTA	1740
Db	1681	GTCTGAGGAGTTACCCATCTAGATATCAAGAGGACTTAAAGAAAGCTGAGCAAAATGTA	1740
Qy	1741	CGACGGTTTATCAGAATTAATCTTCTTAAGGGTGTGATAATTTTCGATATCGCAAGTT	1800
Db	1741	CGACGGTTTATCAGAATTAATCTTCTTAAGGGTGTGATAATTTTCGATATCGCAAGTT	1800
Qy	1801	CAAAGACATGTGCAAGGCTTTAGATGTTAGTCTCTGATGTGGCAGCAGAGTAATCTGTTC	1860
Db	1801	CAAAGACATGTGCAAGGCTTTAGATGTTAGTCTCTGATGTGGCAGCAGAGTAATCTGTTC	1860
Qy	1861	AGTGGCGGAGAAATAGAGCGGTTTAACTCTTACTTTTGTAAAGCCCAACCGAGGAGATGT	1920
Db	1861	AGTGGCGGAGAAATAGAGCGGTTTAACTCTTACTTTTGTAAAGCCCAACCGAGGAGATGT	1920
Qy	1921	GGCTAAAGGCTTTTAAAGACACGCGCTCTGAGCGCGTGTATGTTCTGAAACCGACATCCGA	1980
Db	1921	GGCTAAAGGCTTTTAAAGACACGCGCTCTGAGCGCGTGTATGTTCTGAAACCGACATCCGA	1980
Qy	1981	AGAGGTGAACGTAATAAATTTTCTATGCTGAGAAAGGAGATTTGCTGTGTGTCAGA	2040
Db	1981	AGAGGTGAACGTAATAAATTTTCTATGCTGAGAAAGGAGATTTGCTGTGTGTCAGA	2040
Qy	2041	AGTCAATGTTTGAAGAACTCTAACTTAGACACAGGAGTTGGAGTCCCTCAACGATTT	2100
Db	2041	AGTCAATGTTTGAAGAACTCTAACTTAGACACAGGAGTTGGAGTCCCTCAACGATTT	2100
Qy	2101	CCATAGGCTTTGCTGATAGTGTGATTACAAAGCAAAATGGCATCGGTGTCTACACTGG	2160
Db	2101	CCATAGGCTTTGCTGATAGTGTGATTACAAAGCAAAATGGCATCGGTGTCTACACTGG	2160
Qy	2161	CTCACTCAAAAGTTCAACAAATGAAGAACTATGTGGACAGTTTGGCAGCTTCGTTGTCCG	2220
Db	2161	CTCACTCAAAAGTTCAACAAATGAAGAACTATGTGGACAGTTTGGCAGCTTCGTTGTCCG	2220
Qy	2221	CACGTATCAAAATCTATGCAAGTCACTAAGAGTGAAGTGGGTATGATTTCTGATTTCCAG	2280
Db	2221	CACGTATCAAAATCTATGCAAGTCACTAAGAGTGAAGTGGGTATGATTTCTGATTTCCAG	2280
Qy	2281	GGAGAAAGTTGGTGTGGGATGTCCTTTTGAAGAAAGTGGCTCTCAAACTCGCGGCA	2340
Db	2281	GGAGAAAGTTGGTGTGGGATGTCCTTTTGAAGAAAGTGGCTCTCAAACTCGCGGCA	2340
Qy	2341	AGGTCAATTCATGGGAGTTGCTCTGGATTACAGGGGAAATTTTACTGCACCTTCTATC	2400
Db	2341	AGGTCAATTCATGGGAGTTGCTCTGGATTACAGGGGAAATTTTACTGCACCTTCTATC	2400
Qy	2401	TTATGAAGGAGATAGAAATGTTGATGAGGCACTGAGAGAGGTTGGCTGTATCATCTGA	2460
Db	2401	TTATGAAGGAGATAGAAATGTTGATGAGGCACTGAGAGAGGTTGGCTGTATCATCTGA	2460
Qy	2461	TACAAATGTTATTTCTGATATTCGAAAGCTCCAAATCTCGAGGAAACAAATGAGAGCGG	2520
Db	2461	TACAAATGTTATTTCTGATATTCGAAAGCTCCAAATCTCGAGGAAACAAATGAGAGCGG	2520
Qy	2521	TGAACCCCAACCACTACTGCAAGATGTTACTTTGTGGATGGGTGCTGGTTGTGGAAA	2580
Db	2521	TGAACCCCAACCACTACTGCAAGATGTTACTTTGTGGATGGGTGCTGGTTGTGGAAA	2580
Qy	2581	GTAACAAAGGAGATTTTGAAGATTTGATCTTTGATGAGGATTTGATCTTGGTCTCTGAAA	2640
Db	2581	GTAACAAAGGAGATTTTGAAGATTTGATCTTTGATGAGGATTTGATCTTGGTCTCTGAAA	2640
Qy	2641	ACAAGCTGCTGTATGATCAGAAAGGCTTAATTTCTATCTGGACTGATAGAGCCACAT	2700

Db 2641 ACRAGCTGCTGATGATCAGAAAGGGCTAAATTCATCTGGACTGATAAGAGCCACAAT 2700
Qy 2701 GGACAATGTGAGAACGGTAGATTCACTTCTAATGATCCAAACCGCGATCACACAAGAG 2760
Db 2701 GGACAATGTGAGAACGGTAGATTCACTTCTAATGATCCAAACCGCGATCACACAAGAG 2760
Qy 2761 GCTTTTATGTAGTAAGGGTGTATGCTGCAACCGGTTGTGTTAACTTCTCTGGTCTTAT 2820
Db 2761 GCTTTTATGTAGTAAGGGTGTATGCTGCAACCGGTTGTGTTAACTTCTCTGGTCTTAT 2820
Qy 2821 CTCCTGTTGGACATCGCATACATTTACGGAGATACACAGCAGATTCCTTTCAATTAACAG 2880
Db 2821 CTCCTGTTGGACATCGCATACATTTACGGAGATACACAGCAGATTCCTTTCAATTAACAG 2880
Qy 2881 AGTTCAGAAATTTCCCGTATCCCAACATTTTGAGAAAGCTGCAAGTGGATGAAGTTGAGAT 2940
Db 2881 AGTTCAGAAATTTCCCGTATCCCAACATTTTGAGAAAGCTGCAAGTGGATGAAGTTGAGAT 2940
Qy 2941 GAGGAGGACCACTGAGATGCCAGGTGATGTGAATTTTTTCTTACAAATCGAAGTACGA 3000
Db 2941 GAGGAGGACCACTGAGATGCCAGGTGATGTGAATTTTTTCTTACAAATCGAAGTACGA 3000
Qy 3001 AGGAGCGGTGACAACTTCAACTGTACAAAGCTCGGTCTCATCTGAGATGATAGGCGG 3060
Db 3001 AGGAGCGGTGACAACTTCAACTGTACAAAGCTCGGTCTCATCTGAGATGATAGGCGG 3060
Qy 3061 TAAGGGAGTACTAAACAGTGTTCCAAACCACTAAAGGGGAAATTTGTAACCTTCACTCA 3120
Db 3061 TAAGGGAGTACTAAACAGTGTTCCAAACCACTAAAGGGGAAATTTGTAACCTTCACTCA 3120
Qy 3121 GGCTGATAAATTTGAGTTAGAGGAGAGGGCTATAAGAAATGTGAACCCGTTCAATGAGAT 3180
Db 3121 GGCTGATAAATTTGAGTTAGAGGAGAGGGCTATAAGAAATGTGAACCCGTTCAATGAGAT 3180
Qy 3181 CCAAGGAGAACTTTGAGATGTGCTCGTGTGATGAGTGGAGGCACTCCACTGACTCT 3240
Db 3181 CCAAGGAGAACTTTGAGATGTGCTCGTGTGATGAGTGGAGGCACTCCACTGACTCT 3240
Qy 3241 GATTTCCAAGTCTTCCCGCATGTTCTAGTCCGCTCTGATAGACACACAAAGAGCTTCAA 3300
Db 3241 GATTTCCAAGTCTTCCCGCATGTTCTAGTCCGCTCTGATAGACACACAAAGAGCTTCAA 3300
Qy 3301 ATATTACACCGTAGTTAGATTCCTTTAGTACAGATAATAGTAGTATGTTCTTTAAAG 3360
Db 3301 ATATTACACCGTAGTTAGATTCCTTTAGTACAGATAATAGTAGTATGTTCTTTAAAG 3360
Qy 3361 CTCCTTCTTTAGAAATGTATGTTAGAGCAGGTAGTAGAGCAATTTACAGATGGA 3420
Db 3361 CTCCTTCTTTAGAAATGTATGTTAGAGCAGGTAGTAGAGCAATTTACAGATGGA 3420
Qy 3421 TGCAGTGTCAAAGGTCAATACTCTTTGTGGCAACACCTAAATCAGGAGACTTTCCAGA 3480
Db 3421 TGCAGTGTCAAAGGTCAATACTCTTTGTGGCAACACCTAAATCAGGAGACTTTCCAGA 3480
Qy 3481 TCTACAGTCTATTACGATGATGCTCCCTGGTAAATAGTACTATCTTAAACAGATGA 3540
Db 3481 TCTACAGTCTATTACGATGATGCTCCCTGGTAAATAGTACTATCTTAAACAGATGA 3540
Qy 3541 TGCTGTTACCATGAGGTTACGTGATATAGTCTTAAATGTGAAGGATTTGTTCTTGATTT 3600
Db 3541 TGCTGTTACCATGAGGTTACGTGATATAGTCTTAAATGTGAAGGATTTGTTCTTGATTT 3600
Qy 3601 TTCCAAAGTATTCGGATGCCAAAGGAGGTGAACCCATGCTAGAGCCAGTTTTCGGTAC 3660
Db 3601 TTCCAAAGTATTCGGATGCCAAAGGAGGTGAACCCATGCTAGAGCCAGTTTTCGGTAC 3660
Qy 3661 CGCGCGGACCGCAAGGGTGCAGACTACTCGAAAAATCTGGTTGCAATGATTAAG 3720
Db 3661 CGCGCGGACCGCAAGGGTGCAGACTACTCGAAAAATCTGGTTGCAATGATTAAG 3720
Qy 3721 AAATTTCAACCCACAGACTGAGGGGACGATTTGACATTTGAGAGCACCGCATCTGTTGT 3780
Db 3721 AAATTTCAACCCACAGACTGAGGGGACGATTTGACATTTGAGAGCACCGCATCTGTTGT 3780

Db 3721 AAATTTCAACCCACAGACTGAGCGGGAAGATTTGACATTTGAGAGCACCGCATCTGTTGT 3780
Qy 3781 AGTAGATAAGGTTTTTTTGTATAGCTATTTTATTTAAAAAGAAAAATACACAAAAATATTGC 3840
Db 3781 AGTAGATAAGGTTTTTTTGTATAGCTATTTTATTTAAAAAGAAAAATACACAAAAATATTGC 3840
Qy 3841 TGGAGTGATGACGAAGGATTTCAATGATGAGATGTTTGGAAAAAGAAAGTACTATT 3900
Db 3841 TGGAGTGATGACGAAGGATTTCAATGATGAGATGTTTGGAAAAAGAAAGTACTATT 3900
Qy 3901 GGACGACTTGGCTAACTACAAATTTTACAGATCTCCCGGCCATTCGATCAGTACAGACAT 3960
Db 3901 GGACGACTTGGCTAACTACAAATTTTACAGATCTCCCGGCCATTCGATCAGTACAGACAT 3960
Qy 3961 GATCAAGGCTCAACCAAAAAGAAATTTGACATTTCAATTTAGAAATGTTTGGCCGGTTCTCTGCTCT 4020
Db 3961 GATCAAGGCTCAACCAAAAAGAAATTTGACATTTCAATTTAGAAATGTTTGGCCGGTTCTCTGCTCT 4020
Qy 4021 GCAAAACAATTTGTCTACCATTTGAAAGCAGATCAACGGTATTTTGGCCGGTTCTCTGCTCT 4080
Db 4021 GCAAAACAATTTGTCTACCATTTGAAAGCAGATCAACGGTATTTTGGCCGGTTCTCTGCTCT 4080
Qy 4081 TACAAGGTTGCTGCTCGAGGCAATTTGATTTCTAAGAAAGTCTTTTCTTTTACTAGGAAAC 4140
Db 4081 TACAAGGTTGCTGCTCGAGGCAATTTGATTTCTAAGAAAGTCTTTTCTTTTACTAGGAAAC 4140
Qy 4141 TCCAGAACAGATTTCAAGAAATTTTCTCGGATCTCGACTCGCAGCTTCTATGATGTT 4200
Db 4141 TCCAGAACAGATTTCAAGAAATTTTCTCGGATCTCGACTCGCAGCTTCTATGATGTT 4200
Qy 4201 AGAACTGGATATTTCTAAGTATGATTAAGTCAAGAACGAGTTTCATTTGTGTGAGTA 4260
Db 4201 AGAACTGGATATTTCTAAGTATGATTAAGTCAAGAACGAGTTTCATTTGTGTGAGTA 4260
Qy 4261 TGAATAATGAAAGATTTGGGTCTCAATGAGTTTGGCCGAAAGTGTGGAACCAAGGCA 4320
Db 4261 TGAATAATGAAAGATTTGGGTCTCAATGAGTTTGGCCGAAAGTGTGGAACCAAGGCA 4320
Qy 4321 CAGAAAAACAATTTGAAGGATTTACATTTGCTGGAATCAAGACATGCTGTGGTATCAAAG 4380
Db 4321 CAGAAAAACAATTTGAAGGATTTACATTTGCTGGAATCAAGACATGCTGTGGTATCAAAG 4380
Qy 4381 GAAAAGCGGTGATGCTACTTTTCATCGGCAATTAAGTATTAATAGCAGCTTGTGGG 4440
Db 4381 GAAAAGCGGTGATGCTACTTTTCATCGGCAATTAAGTATTAATAGCAGCTTGTGGG 4440
Qy 4441 TTCAAATGTTACCGATGAAAGGTCATAAAGGTCCTTTTGTGGAGACGATTTCCGTTT 4500
Db 4441 TTCAAATGTTACCGATGAAAGGTCATAAAGGTCCTTTTGTGGAGACGATTTCCGTTT 4500
Qy 4501 GTATTTTCCAAAGGTTTGGATTTCCCTGACATTTCAAGTCAATGTGCTAAATCTCATGTGAA 4560
Db 4501 GTATTTTCCAAAGGTTTGGATTTCCCTGACATTTCAAGTCAATGTGCTAAATCTCATGTGAA 4560
Qy 4561 TTTTGGGCCAAACTGTATAGAAAGGTACCGTTACTTTTGTGGTAGATACATCATACA 4620
Db 4561 TTTTGGGCCAAACTGTATAGAAAGGTACCGTTACTTTTGTGGTAGATACATCATACA 4620
Qy 4621 CCATGATAAGGGAGCAATAGTGTATGATCTCTTTGAAGTGTGATCTTCCAAACTTGGGG 4680
Db 4621 CCATGATAAGGGAGCAATAGTGTATGATCTCTTTGAAGTGTGATCTTCCAAACTTGGGG 4680
Qy 4681 AAAACATATCAAGGATTTGATCACTTTAGAAAGTAAAGGGTGTCTTTTGTGCGATGTGC 4740
Db 4681 AAAACATATCAAGGATTTGATCACTTTAGAAAGTAAAGGGTGTCTTTTGTGCGATGTGC 4740
Qy 4741 TTGTTGCTCGGAAACTGTTGCTTAGGCTTTCCGAGCTGAAACGAGCTATCAAGGAGGT 4800
Db 4741 TTGTTGCTCGGAAACTGTTGCTTAGGCTTTCCGAGCTGAAACGAGCTATCAAGGAGGT 4800
Qy 4801 TCATAAAAACCGCATTTGATGTTGCTTTTAAATTTGTTAAACAAATTTTGTGTGA 4860
Db 4801 TCATAAAAACCGCATTTGATGTTGCTTTTAAATTTGTTAAACAAATTTTGTGTGA 4860

```
Qy 4861 TAAATTTTATTTAGAACTTTGTTTTTAAATGGCTGTGTCTCTCAGAGATACTCTCAAAA 4920
Db 4861 TAAATTTTATTTAGAACTTTGTTTTTAAATGGCTGTGTGTCTCAGAGATACTCTCAAAA 4920
Qy 4921 TTACCGAGTTTCATTGATCTTTTCGAAACAGGATGAGATACTTCCGGCCATTCAATGACTAAGG 4980
Db 4921 TTACCGAGTTTCATTGATCTTTTCGAAACAGGATGAGATACTTCCGGCCATTCAATGACTAAGG 4980
Qy 4981 TCAAGAGTTGTAGAAATATCGACTGTGGACAAAGATATATGGCTGTGTTAAGAAATGATAGTCTTT 5040
Db 4981 TCAAGAGTTGTAGAAATATCGACTGTGGACAAAGATATATGGCTGTGTTAAGAAATGATAGTCTTT 5040
Qy 5041 CTGATGTAGATTTTACTTTAAAGGTGTAAAGTTAGTTTAAAGAAAGGTATGTGTGCTTAGCTG 5100
Db 5041 CTGATGTAGATTTTACTTTAAAGGTGTAAAGTTAGTTTAAAGAAAGGTATGTGTGCTTAGCTG 5100
Qy 5101 ATTTGGTAGTGTCTGGGGAGTGGGAATCTCCCGGATAACTCCGCTGGTGGTGTCAAGTGTGTT 5160
Db 5101 ATTTGGTAGTGTCTGGGGAGTGGGAATCTCCCGGATAACTCCGCTGGTGGTGTCAAGTGTGTT 5160
Qy 5161 GTATTGTAGATPAAGAAATGAAAGAGTAAAGGAAGCAACGCTGGGTGCGTATCACGCC 5220
Db 5161 GTATTGTAGATPAAGAAATGAAAGAGTAAAGGAAGCAACGCTGGGTGCGTATCACGCC 5220
Qy 5221 CTGCTTCGAAAAAGAAATTTCTTTTAAAGCTTAATCCCTAATTTTCAATACATCCGAGG 5280
Db 5221 CTGCTTCGAAAAAGAAATTTCTTTTAAAGCTTAATCCCTAATTTTCAATACATCCGAGG 5280
Qy 5281 ATGCTGAGAGACCCCGTGGCAAGTGTAGTGAATATCAAAAGGAGTGGCTATGGAAGAG 5340
Db 5281 ATGCTGAGAGACCCCGTGGCAAGTGTAGTGAATATCAAAAGGAGTGGCTATGGAAGAG 5340
Qy 5341 GATACGTCTCTTTATCTTTGGAGTTCGTTTTCAATTTGTGTAGTACATAAAAAATTAATGTA 5400
Db 5341 GATACGTCTCTTTATCTTTGGAGTTCGTTTTCAATTTGTGTAGTACATAAAAAATTAATGTA 5400
Qy 5401 GAAAGGTTTGAGGGAACGTATTTTGTAGTGTGACAGAGCGGTCCGCCAATTGAACTCACTG 5460
Db 5401 GAAAGGTTTGAGGGAACGTATTTTGTAGTGTGACAGAGCGGTCCGCCAATTGAACTCACTG 5460
Qy 5461 AAAAGGTTTGTGAGGAGTTCGTGTGATGAAGTACCAGTGGCTGCAAACTCGAAAAGGTTTC 5520
Db 5461 AAAAGGTTTGTGAGGAGTTCGTGTGATGAAGTACCAGTGGCTGCAAACTCGAAAAGGTTTC 5520
Qy 5521 CGGAAAAACAAAAAGAAATGCTAGTAAATATGTTAAATAAAGAAATAAATAACAGTG 5580
Db 5521 CGGAAAAACAAAAAGAAATGCTAGTAAATATGTTAAATAAAGAAATAAATAACAGTG 5580
Qy 5581 GTAAGAGGGTTTTAAATTCAGGAAATGAGGATAATGTTAAGTGATGACAGGCTCTATCG 5640
Db 5581 GTAAGAGGGTTTTAAATTCAGGAAATGAGGATAATGTTAAGTGATGACAGGCTCTATCG 5640
Qy 5641 CGTCATCGAGTACGTTTTTAATCAATATGCTTTATACATCAACTCTCCGACCAATTTGT 5700
Db 5641 CGTCATCGAGTACGTTTTTAATCAATATGCTTTATACATCAACTCTCCGACCAATTTGT 5700
Qy 5701 TTACTTATCTTCCGCTTACGAGATCTCTGTGCAGCTGATCAATCTGTGTACAAATGCAAT 5760
Db 5701 TTACTTATCTTCCGCTTACGAGATCTCTGTGCAGCTGATCAATCTGTGTACAAATGCAAT 5760
Qy 5761 GGGTAAACAGTTTCAACGCAACAGCTAGGACCAACAGTCCAAACAGCAATTTGCGGATGC 5820
Db 5761 GGGTAAACAGTTTCAACGCAACAGCTAGGACCAACAGTCCAAACAGCAATTTGCGGATGC 5820
Qy 5821 CTGGAACCTGTGCTAGTATGACAGTGGAGATTTTCCCTGCATCGGATTTCTATGTGTATG 5880
Db 5821 CTGGAACCTGTGCTAGTATGACAGTGGAGATTTTCCCTGCATCGGATTTCTATGTGTATG 5880
Qy 5881 ATATAATTCGACGCTTGTATCGGTTGATCAACGCGGTTTAAATAAGCTTTGATCTAGAAA 5940
Db 5881 ATATAATTCGACGCTTGTATCGGTTGATCAACGCGGTTTAAATAAGCTTTGATCTAGAAA 5940
```

```
Qy 5941 TAGAATAATAGAGTTTGATAATCAACCCGCAACCGAATACTACTGAAATCGTTAACGGAC 6000
Db 5941 TAGAATAATAGAGTTTGATAATCAACCCGCAACCGAATACTACTGAAATCGTTAACGGAC 6000
Qy 6001 TCAGAGGGTAGACGATCTACTGTAGCTATAAAGGGCTTCAATCAATAATTTGGCTAATGA 6060
Db 6001 TCAGAGGGTAGACGATCTACTGTAGCTATAAAGGGCTTCAATCAATAATTTGGCTAATGA 6060
Qy 6061 ACTGGTTTCGTGGAACTGGCATGTTCAATCAAGCAGGCTTTGAGACTGCTAGTGACATTGT 6120
Db 6061 ACTGGTTTCGTGGAACTGGCATGTTCAATCAAGCAGGCTTTGAGACTGCTAGTGACATTGT 6120
Qy 6121 CTGGACCACAACTCCGGCTACTTAGCTATTGTTGTGAGATTTCCTAAAAATAAAGTCGCTG 6180
Db 6121 CTGGACCACAACTCCGGCTACTTAGCTATTGTTGTGAGATTTCCTAAAAATAAAGTCGCTG 6180
Qy 6181 AAGACTTTAAAAATTCAGGGTGGTGATACCAAAATCAGCAGTGGTTGTTCCGTCACCTTAAA 6240
Db 6181 AAGACTTTAAAAATTCAGGGTGGTGATACCAAAATCAGCAGTGGTTGTTCCGTCACCTTAAA 6240
Qy 6241 TATAACGATTGTCTATCTCGATCCCAACAGTTAAACCATGTGATGGTGATCTACTGTGTA 6300
Db 6241 TATAACGATTGTCTATCTCGATCCCAACAGTTAAACCATGTGATGGTGATCTACTGTGTA 6300
Qy 6301 TGGCGTAAAAATTCAGGAGAGTTCGAAATCTCCCTAACCCCGGTAGCGGCCCA 6355
Db 6301 TGGCGTAAAAATTCAGGAGAGTTCGAAATCTCCCTAACCCCGGTAGCGGCCCA 6355

RESULT 4
AB078435
LOCUS
DEFINITION
Tobacco mild green mosaic virus complete genome, strain:Japanese.
ACCESSION
AB078435
VERSION
AB078435.1
KEYWORDS
GI:18253266
SOURCE
Tobacco mild green mosaic virus
ORGANISM
Tobacco mild green mosaic virus
REFERENCE
1
Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
AUTHORS
Okuno, T., Hamada, H., Takeuchi, S., Morishima, N., Yoshimoto, E. and
Hikichi, Y.
TITLE
Nucleotide sequence of the Japanese isolate of Tobacco mild green
mosaic virus
JOURNAL
Unpublished
REFERENCE
2 (bases 1 to 6356)
AUTHORS
Okuno, T., Hamada, H., Takeuchi, S., Morishima, N., Yoshimoto, E. and
Hikichi, Y.
TITLE
Direct Submission
JOURNAL
Submitted (18-JAN-2002) Tetsuro Okuno, Kyoto University;
Kitashirakawa Oiwakecho, Sakyo 606-8502, Japan
(S-mail:okuno@kais.kyoto-u.ac.jp. Tel:81-75-753-6131,
Fax:81-75-753-6131)
FEATURES
Location/Qualifiers
source
1..6356
/organism="Tobacco mild green mosaic virus"
/virus
/mol_type="genomic RNA"
/strain="Japanese"
/db_xref="taxon:12241"
/notes="RNA transcripts synthesized in vitro from the cDNA
clone of TMGMV-J direct systemic infection in Nicotiana
benthiana-TMGV-Japanese strain"
72..4901
/notes="position 3405..3407 is readthrough product"
/codon_start=1
/transl_except=(pos:3405..3407,aa:OTHER)
/product="183kDa protein"
/protein_id="BAB93986.1"
/db_xref="GI:18253267"
/translacion="MAHIQSTISNALLSVSGKNTLVNDLARRMYDTAVDFNARDR
RPKNVPKTISEBQTLVATYREFQITFYNTONAVHSLAGLPALSLVLMLOPVG
SPYTDIGNFAHLFKGRDVFHCCMFLDIRINRHEGQKDSIEMYLRLSRSNKVIP
```

EFORAAFNRYAEAPNEVCSKTFQDCHHPENSGRRVAISLHSLYDI PVHREGAALI
SKOJICIAAFSEALLQDTEVLNIGATFKREGDDVVSFFPADESTLNNHKTGN
LIHYCVSPFASRIVYKPEFLVTRVNTFKCTVDYILIKSVQVCGDSQRYE
AMEDATFYKALAMFENTERAILEDYAPVPMKMOXIVVPJFESITKOTSESR
IVNRDVFVNLHARTYQAKALYQNVLSFVRESRVRVINGVARGSEMWDKXALIQP
LSMTFLOTUKALLODIIWVGKFRCLDKTSTLWDEWQVFGNVPYTI KERLVSRK
LYDSENALIKIIPDLVYQNRVPAEYVTSKSELEPHDIKKDLAEACOMYDALSELSIL
RDADNFIDIAKDMKXALCVNDVAARVIVAVENRSGLLITFDKPTTEENVAKALKST
ASEAVKLEPSTDEWVNVQVQKMSVTAERGKQVCAEHSGLTNANLQELUESLNDPHKACV
DSVITIKOMESVYTSGLVQVQKMYNVDSLAASLQATVSNLCSKLQVGYDSDSRKACV
GMYDITLKWLLKPAAGHSWGVYDYGKMTALLSYEGEDMYVTESDMRWVAVSSDT
WYVSOTAKLONKTRWDEGHEPTEAKVMLVGVQCGCKTEKILERLVDDEDLIVPG
KQAAAMIRREANSGLIRATMDNVYDSDFLMHPKPRTHKELRDEGLMLTCTGVNFL
LWASGDLAYIVGDTQOIPINRVQNPYVFKPKFKLQDVEVEMWRTMLTCEPQGVNPL
QSKTEGAVSTISYQKRSVASEMI GKGVLNVSVKPKGKIVFTQADKFEKEBKYN
VNTVHEIQTGFEDSVLEWYVATPLTILSKSPSVNVALTRTKTSFKYITVLVDPLVQ
IISDLSGLSELMYVWBAQRQLOMDAVFGKHNIVPATPKSGDPDQIVLQVCL
PGNSTILNKYDAMTWMLRNDGSDVDSFKSIIPMKVPEKCEPLVKTAAEPKTV
AGLLENVAMIKNFNAVPLDTGIDESTVASVDVDPKDSVFKKCEPLVKTAAEPKTV
DSMWRLWQKQERSTIGOLANYFVLDPAIDOKYKMIKAQPKQLDLSIQNEYPALOTI
VYHSKOINGIFGPGVSELETRLLLEAVDSQKELFTFRKTPQIQEFPDSDLSHVPMVLD
ELDISIKSGNEBFCABVSEIWKRLGJNLMPEKVIKGAQPGDSDVYFPGKGLDPTQSCAN
QRKSDVFTYPIGNVTIJAALGSLMIPMEKVIKGAQPGDSDVYFPGKGLDPTQSCAN
LWMNFVEAKYKKGYFCGCRYIITHDGKAIIVYDPLKILSKLGAHITKDYHLEELRV
SLDVCASGLINCAYFPQLNAAIKVHVKHTAIDGSAFNCVNCWFLCDKFLFTPLTFLNGC"

5

```

4. ....350)
/codon_start=1
/product="126kDa protein"
/protein_id="BAB83987.1"
/db_xref="GI:18253268"
/tranlation="MAHISQTNALLSEVSQKNTLVNDLARRMYDTVADEENARDR
RPKVNPPTISEQTLTLVNAIPETQITFYNTQNAHSLAGLRALLEYMLQVPGY
EYFDITGGNFALPKFRGYDHCMPENLDIDIRHGGKQXDSIMLYLSLHNSKVP
EFOFAAFNPYAEAPNEVSCSKTFCQIRHPDENGRRYVASHLSYDIPYHFEAALI
SKDITCIYAAPHSEALLDQTEVTLNIGATFKREGDDVSEFFADESTLNSHXYKN
ILHVVXSYPPASSRIYVPEKLEIVRVNTWCCKPTKDVTYILKXSVQGVCSNDQYVE
AMEDATFYKALAMFNTERRAILRDTSANVFWPEPKMDWIVLPFEGSITSKMTREY
LVNFTDITVNHIRTYQAKULTQNVLSYRSRSRVINGVTVARSEWDVDRKAILOP
1SWMTFQTLKLAALQDDIYMGKFRCLDKTSGELIWDVEGYPFGNVTPTIKERLVRSKI
LIDVSENALKIPLDYVTKWDRFAVYVIAHLNDRITDKOLDEAALDLSGLSL
RDADNP1AKFKCMKALCVNPQDAVRYVAENRSGLTITLPOKPEENVAKALSTL
ASEAVVLEPTSDVGMVKFESAERGLPCVAESHGLTNANLHQELESVDHPKACV
DSVITQKMASVVTYSGLSKQVMQNVYDLSLAASLANTSVNLSKSLKVDVSDSRKV
GMDVTIKWLLKPAKAGWYVDIYKGMFTALLSYEGDMVSDVWDWRVASVSDT
WVSDTAKLQNKRTWRDEPNTAKWLVDPGVGPKCTKEILERVLDDLEDILVPG
KQAAAMIRRNANSGSLIRATMDNVTWDSFLMHPKRTYKHALFTIDEGMLHTCGVNF
VLISGCDIAVITYGDTOQIPFINRVQNPYVPKHFELQVDEVERMTTTLRCPGDWNVFL
QSKYEGAVSTTSDVSVSSESMIGKGVLSNVSPVLEKGIIVTPTQADKPELEBKQYV
NVNTHETIQGSTFEDVSLVRLATATPITLISKSGSPHVLVAULTHTTKSYFKYITVLVDPLVQ
1ISDLSLSFLLEMTYVFAAGR"
4891. ....5661

```

CDS

```

/codon_start=1
/product="28kDa protein"
/protein_id="BAB83988.1"
/db_xref="GI:18253269"
/translations="MAVSLRDVTVKISEFDLSKODEILPAPMTKVGSRVISTVDKIMAK
VKNDLSVDPLLAGKVLKNGYVSLAGIYVSGEKNLPDNCRGGVSVCIQDKMRKSRKE
ALTGAYHAPACKNGKFEKLLPNYSITSEDAKHPQVULVNIKGVAMEEGYCPKLSLFP
SICVWHKQNVKGRGLRRIILRVTDGPIELTEUKVEEFDVDFVPMVKLERCPKTKKRVV
GNSYNNKKINNSGKKGLKVEEIEDSDDESSTASST"
5667. 6146
/codon_start=1
/product="17kDa protein"
/protein_id="BAB83989.1"
/db_xref="GI:18253270"
/translations="MPYTNISPSQFYVLLSSAYADPVQLINI.CTNALGNQFQOQARTT
VOQOQADAKPESIVITRVPASDFVYRNTSLDPLITALINSFDTRNRIIEVDNOPA
PNTTEIYNATORVDDATVAIASINMIANEIVRGCTGMPNOAGPPTASGIVVTTTTPAT"

```

CDS

```

/codon_start=1
/product="17kDa protein"
/protein_id="BAB81989.1"
/db_xref="GI:18253270"
/translations="MPTYNSSQFVYLSSAYADVQLINLTNALGNQFQQTQRT
VQOGEADAKWPSITVRPPEADQVYRNTSLDPLITALINSFDTRNRIIEVDNQA
VQOQELUNATQVDDADATVIRASINNIANETVRCGTGMPNOAGPPTASGIWTTTTPAT"

```

ORIGIN

Query Match 93.7%; Score 5951.6; DB 14; Length 6356;
Best Local Similarity 96.4%; Pred. No. 0;
Matches 6131; Conservative 0; Mismatches 219; Indels 8; Gaps 4;

Qy	2	ATGTTTAAATAGT	TTTCGACACACAA	CAATTA	TA	AAACAA	AAACAA	CAATAT	TACAAACA	CAAA	61
Db	3	ATGTTTAAATAGT	TTTCGACACACAA	CAATTA	TA	AAACAA	AAACAA	CAATAT	TACAAACA	CAAA	62
Qy	62	ACAACAAATAGT	GCACACATACA	CAATCT	TA	TAATTTAG	CAACGCCCTT	CTTCTG	AAACGCGTGAGT	121	
Db	63	ACAACAAATAGT	GCACACATACA	CAATCT	TA	TAATTTAG	CAACGCCCTT	CTTCTG	AAACGCGTGAGT	122	
Qy	122	GGTAAACACACT	CTCGTTAATGAC	CTTTGCAAG	AGGCGCAT	GTACGAT	ACGGCCGTGGAA	181			
Db	123	GGTAAACACACT	CTCGTTAATGAC	CTTTGCAAG	AGGCGCAT	GTACGAT	ACGGCCGTGGAC	182			
Qy	182	GAATTTAAACCG	CGCGAGCGTAG	ACCACAA	AGGTCAA	CTTTTCC	AAAACTATTACG	GAAGAG	241		
Db	183	GAATTTAAACCG	CGCGAGCGTAG	ACCACAA	AGGTCAA	CTTTTCC	AAAACTATTACG	GAAGAA	242		
Qy	242	CAAAACGCTTCT	AGTCTCTCAAC	CGGTAC	CCCGAGTCC	AGATTAC	CTTTTATA	TACTCAA	301		
Db	243	CAAAACGCTTCT	AGTCTCAAC	CGGTAC	CCCGAGTCC	AGATTAC	CTTTTATA	TACTCAA	302		
Qy	302	AATGCCGTGACA	AGTTTGGCTGAG	AGTTTGAG	AGCAAT	TAGAA	TTTGGAA	TATCTGATGCTA	361		
Db	303	AATGCCGTGACA	AGTTTGGCTGCG	TTTGAG	AGCAAT	TAGAA	TTTGGAA	TATCTGATGCTA	362		
Qy	362	CAAGTTCCCTAT	CGATCGCCGAC	ATATGAT	TATAGT	TGGGA	CACTTTG	CAGCACATTTTGTTC	421		
Db	363	CAAGTTCCCTAT	CGATCA	CCGACAT	TATATAGT	TGGGA	CACTTTG	CAGCACATTTTGTTC	422		
Qy	422	AAAGGCAAGGAT	TAACGTGCA	TTGCTGTAT	GCCCAATCT	TGGACAT	TACGAGAT	TAATGAGG	481		
Db	423	AAAGGCAAGGAT	TAACGTGCA	TTGCTGTAT	GCCCAATCT	TGGACAT	TACGAGAT	TAATGAGG	482		
Qy	482	CAGAGGACAA	AGAGCTCAAT	TGAGATG	TATTTG	TCCAGAT	TGTCTCGTCT	CTAACAAG	541		
Db	483	CAGAGGACAA	AGAGACTCAAT	TGAGATG	TATTTG	TCCAGAT	TGTCTCGT	CTAACAAG	542		
Qy	542	GTAATTCCTGAG	TTTCAAAGGGAG	GGCTTTTAA	CAGGTTATG	CAGAAG	CTCCCA	CGAAGTC	601		
Db	543	GTAATTCCTGAG	TTTCAAAGGGCG	GGCTTTTAA	TAGATATG	CAGAAG	CTCCGA	ACGAAGTC	602		
Qy	602	TGCTGCTCTAAA	ACTTTTCAGGAT	TGCGNAT	ATACATCC	CGCCAGAG	ATAGTGTG	TAGNAGA	661		
Db	603	TGCTGCTCTAAA	ACTTTTCAGAG	ATGTCGA	ATATCC	ACAGAG	ATAGTGTG	TAGNAGA	662		
Qy	662	TAGCTGTTCCT	CTGCACAGTT	TGTATGAT	TATTTCT	GTCAT	GAGTTTGAG	CTGGTTA	721		
Db	663	TAGCTATTTCT	CTGCACAGTT	TGTATGAT	TATTTCT	GTCAT	GAGTTTGAG	CTGGTTA	722		
Qy	722	ATATCTAAGA	AAATATACAT	GTATGTATG	AGCTTCCA	ATTTTGG	CAGAGCA	TATTTACTTA	781		
Db	723	ATATCTAAGG	ACATACATAT	GTATGTATG	AGCCCTTCCA	ATTTTCT	GAGCA	TATTTACTTA	782		
Qy	782	GACACAGCGAG	GTACGCTTAAT	GAAATAG	GAAGCGCA	CTTTTCAA	AAAGAGAG	GTGATGAT	841		
Db	783	GACACAGCGAG	GTATCGCTTAAT	GAAATAG	GAAGCGCA	CTTTTCAA	AAAGAGAG	GTGATGAT	842		
Qy	842	GTTTCTTTTCT	TTTGTGATGA	AGACTTTTAA	ATTTATAG	TACTATA	ATACAAAA	TATC	901		
Db	843	GTTTCTTTTCT	TTTGTGATGA	AGACTTTTAA	ATTTATAG	TACTATA	ATACAAAA	TATTT	902		
Qy	902	TTGCATATATG	ATAGTAAATCT	TTTACTTTCT	GCTTCTAG	TAGATAG	TATTTACT	TTTAGGAA	961		
Db	903	TTGCATATATG	ATAGTAAATCT	TTTACTTTCT	GCTTCTAG	TAGATAG	TATTTACT	TTTAGGAA	962		
Qy	962	TTTTTTAGTCT	ACTAGGGTTA	TACTTGGT	TTTGTAA	ATTTTAC	CAAAAG	TAGATCCTATATTT	1021		
Db	963	TTTTTTAGTCT	ACTAGGGTTA	TACTTGGT	TTTGTAA	ATTTTAC	CAAAAG	TAGATCCTATATTT	1022		
Qy	1022	CTGTACAAG	AGTGTATAG	CAAGTAG	GGTGTATG	TATGAT	CAGTTCT	TATGAGGCGATGGAA	1081		
Db	1023	CTGTACAAG	AGTGTATAG	CAAGTAG	GGTGTATG	TATGAT	CAGTTCT	TATGAGGCGATGGAG	1082		

Qy 1082 GACGCTTTGCTTACAGAAACCTTGGCCATGTTCAACACTGAAGAGCAATCTTTAGA 1141
Db 1083 GACGCTTTACTTATAGAAAGCCTTGGCTATGTTCAACACTGAAGAGCAATCTTTAGA 1142
Qy 1142 GACACGGCTTCGGTTAACTTTTGGTCCCTTAAGATGAAGGACATGGTGATAGTACCCTG 1201
Db 1143 GACACGGCTTCGGTTAACTTTTGGTCCCTTAAGATGAAGGACATGGTGATAGTACCCTG 1202
Qy 1202 TTTGAGGGTTCTATTTACCAGCAAAAGATGACAGGAGTGAGGTCATTTGTTAAATCGTGAC 1261
Db 1203 TTCGAGGGTTCTATTTACCAGCAAAAGATGACAGGAGTGAGGTCATTTGTTAAATCGTGAC 1262
Qy 1262 TTTGAGGGTTCTATTTACCAGCAAAAGATGACAGGAGTGAGGTCATTTGTTAAATCGTGAC 1321
Db 1263 TTTGAGGGTTCTATTTACCAGCAAAAGATGACAGGAGTGAGGTCATTTGTTAAATCGTGAC 1322
Qy 1322 AACGTATTTATCTTTGCGGAGTCTATAAGATCCCGGTGATATAATCAATGTTGTTACGTCT 1381
Db 1323 AACGTATTTATCTTTGCGGAGTCTATAAGATCCCGGTGATATAATCAATGTTGTTACGTCT 1382
Qy 1382 AGGTCTGAATGGGATGATAGTAAGCAATCTTCAACCCCTTGTCAATGACTTCTTCTTGG 1441
Db 1383 AGGTCTGAATGGGATGATAGTAAGCAATCTTCAACCCCTTGTCAATGACTTCTTCTTGG 1442
Qy 1442 CAGACTAAGCTGGCTGGCTTCAAGACGATATAGTAATGGAAAGTTTCCGTGCTTGGAT 1501
Db 1443 CAGACTAAGCTGGCTGGCTTCAAGACGATATAGTAATGGAAAGTTTCCGTGCTTGGAT 1502
Qy 1502 AAGACCACTTCTGAACTTATTTGGGATGAGTGGGCAAAATTTTGGAAACGTTTCCCTC 1561
Db 1503 AAGACCACTTCTGAACTTATTTGGGATGAGTGGGCAAAATTTTGGAAACGTTTCCCTC 1562
Qy 1562 ACTATCAAGAGAGATTTGGTGAGCAGGAAATTTCTGATGTAGTGAGATGCTCTGAA 1621
Db 1563 ACTATCAAGAGAGATTTGGTGAGCAGGAAATTTTGGATGTAGTGAGATGCTCTGAA 1622
Qy 1622 ATCAAGATCCAGATCTGTATGTCAATGGAAGACAGGTTCTGATGTAGTAATCAACCAAG 1681
Db 1623 ATCAAGATCCAGATCTGTATGTCAATGGAAGACAGGTTCTGATGTAGTAATCAACCAAG 1682
Qy 1682 TCTGAGGAGTTACCGCATCTGGATATCAAGAAGGACTTGAAGAAGCTGAGCAATGTAC 1741
Db 1683 TCTGAGGAGTTACCGCATCTGGATATCAAGAAGGACTTGAAGAAGCTGAGCAATGTAC 1742
Qy 1742 GACGCTTATCAGAAATTTATCTATCTTAAAGGCTGATATAATTTTCGATATCGGAAATTC 1801
Db 1743 GACGCTTATCAGAAATTTATCTATCTTAAAGGCTGATATAATTTTCGATATCGGAAATTC 1802
Qy 1802 AAAGACATGTGCAAGGCTTTAGATGTTAGTCTGATGTGGCAGCACGAGTAATCGTTGCA 1861
Db 1803 AAAGACATGTGCAAGGCTTTAGATGTTAGTCTGATGTGGCAGCACGAGTAATCGTTGCA 1862
Qy 1862 GTGCGGAGAAATAGAAGCGGTTTAACTCTTACTTTTGAATGAAGCAACCGAGGAAATGTG 1921
Db 1863 GTGCGGAGAAATAGAAGCGGTTTAACTCTTACTTTTGAATGAAGCAACCGAGGAAATGTG 1922
Qy 1922 GCTAAGGCTCTTAAAGCAGCGGCTGTAGGCGCTGTATGTCTTGAACCGACATCCGAA 1981
Db 1923 GCTAAGGCTCTTAAAGCAGCGGCTGTAGGCGGCTGTATGTCTTGAACCGACATCCGAT 1982
Qy 1982 GAGGTGAACGTAAATAAATTTCTATTGCTGAGAAAGGAGATGCTGCTGTGTGACAGAA 2041
Db 1983 GAGGTGAACGTAAATAAATTTCTATTGCTGAGAGAGGAAATTTGCCCGTGTGTGCTGAA 2042
Qy 2042 AGTCATGTTTGAAGGATGCTAACTTAGAGCACAGGAGTTGGAGTCCCTCAACGATTTTC 2101
Db 2043 AGTCATGTTTGAAGGATGCTAACTTAGAGCACAGGAGTTGGAGTCTCTCAACGATTTTC 2102
Qy 2102 CATAAAGCTTCGTTGGATAGTGATTAACAAGCAAAATGGCATCGGTTGTCTACATGGC 2161
Db 2103 CATAAAGCTTCGTTGGATAGTGATTAACAAGCAAAATGGCATCGGTTGTCTACATGGC 2162
Qy 2162 TCACTCAAAGTTCAACAAATGAAGAACTATGTGACAGTGTGGCAGCTTCGTTGTCGCGC 2221

Db 2163 TCACTCAAAGTTCAACAAATGAAGAACTATGTGACAGTGTGGCAGCTTCGTTGTCGCGC 2222
Qy 2222 ACTGTATCAAATCTATGCAAGTCACTAAAGGATGAAGTCCGGTATGATTTCTGATTCAGG 2281
Db 2223 ACTGTATCAAATCTATGCAAGTCACTAAAGGATGAAGTTCGGTATGATTTCTGATTCAGG 2282
Qy 2282 GAGAAAGTTCGGTTCGGTATGTCACCTTGAAGAAAGTGGCTCCTCAAAACCTTCGGCCAAA 2341
Db 2283 GAGAAAGTTCGGTTCGGTATGTCACCTTGAAGAAAGTGGCTCCTTAAACCGCGCGCAAAA 2342
Qy 2342 GGTCAATTCATGGGAGTTGCTCGGATTTACAAGGGGAAAAATGTTTACTGCACTTCTATCT 2401
Db 2343 GGTCAATTCATGGGAGTTGCTCGGATTTACAAGGGGAAAAATGTTTACTGCACTTCTATCT 2402
Qy 2402 TATGAAGGAGATAGAATGGTGACTGAGAGGCACTGAGAGGAGGTGGCTGTATCATCTGAT 2461
Db 2403 TATGAAGGAGATAGAATGGTGACTGAGAGGCACTGAGAGGAGGTGGCTGTATCATCTGAT 2462
Qy 2462 ACAATGATATTTCTGATATTTGCAAAAGCTCCAAAATCTGAGGAAAAACAATGAGAGAGCGT 2521
Db 2463 ACAATGATATTTCTGATATTTGCAAAAGCTCCAAAATCTGAGGAAAAACAATGAGAGAGCGT 2522
Qy 2522 GAAACCCACCAACCTACTCTGCAAAAGATGGTACTTGTGATGGGCTCCCTGGTTGTGGAAG 2581
Db 2523 GAAACCCACCAACCTACTCTGCAAAAGATGGTACTTGTGATGGGCTCCCTGGTTGTGGAAG 2582
Qy 2582 TACAAGGAGATTTTGAAGATTTGATCTTTGATGAGGATTTGATCTTTGGTTCCTGGAANA 2641
Db 2583 TACAAGGAGATTTTGAAGATTTGATCTTTGATGAGGATTTGATCTTTGGTTCCTGGAANA 2642
Qy 2642 CAAGCTGCTGATGATCAAGAAAGGCTAAATCTCTGAGCTGATAAGAGCCACCAATG 2701
Db 2643 CAAGCTGCTGATGATCAAGAAAGGCTAAATCTCTGAGCTGATAAGAGCCACCAATG 2702
Qy 2702 GACAAATGTGAGAAAGCTGATGATTTCACTTTCAATGATCCAAAACCGCCATCACACAAGAG 2761
Db 2703 GACAAATGTGAGAAAGCTGATGATTTCACTTTCAATGATCCAAAACCGCCAAACGCAAGAG 2762
Qy 2762 CTTTTTATGATGAAGGTTGATGTGCAACACCGGTTGTGTTAACTTTCTGGTCTTATC 2821
Db 2763 CTTTTTATGATGAAGGTTGATGTGCAACACCGGTTGTGTTAACTTTCTGGTCTTATC 2822
Qy 2822 TCTGGTTGCGACATCCGATACATTTACGGAGATACACAGCAGATTCCTTTTCATTAACAGA 2881
Db 2823 TCTGGTTGCGACATCCGATACATTTACGGAGATACACAGCAGATTCCTTTTCATTAACAGA 2882
Qy 2882 GTTCAGAAATTTCCCGTATCCCAACATTTTGAGAGCTGCAAGTGAAGTTGAGATG 2941
Db 2883 GTTCAGAAATTTCCCGTATCCCAACATTTTGAGAGCTGCAAGTGAAGTTGAGATG 2942
Qy 2942 AGGAGGACCACTGAGATGCCCAGGTGATGTGAATTTTTTCTTACAAATCGAAAGTACGAA 3001
Db 2943 AGGAGGACCACTGAGGTGCCCAGGTGATGTGAATTTTTTCTTACAAATCGAAAGTACGAA 3002
Qy 3002 GAGCGGTGACACCACTTCAACTGTACACGATCGGTCTCATCTGAGATGATAGCGGT 3061
Db 3003 GAGCGGTGACACCACTTCAACTGTACACGATCGGTCTCATCTGAGATGATAGCGGT 3062
Qy 3062 AAGGGAGTACTAAACAGTGTTCCAACCACTAAAGGGGAAAAATTTGTAACCTTTCACTCAG 3121
Db 3063 AAGGGAGTACTAAACAGTGTTCCAACCACTAAAGGGGAAAAATTTGTAACCTTTCACTCAG 3122
Qy 3122 GCTGTATAAATTTGAGTTAGAGGAGAGGCTTATAAGAAATGTGAACCGGTTTCATGAGATC 3181
Db 3123 GCTGTATAAATTTGAGTTAGAGGAGAGGCTTATAAGAAATGTGAACCGGTTTCATGAGATC 3182
Qy 3182 CAAGGAGAAACCTTTGAAGATGTGCTGGTCAAGATTTGACGGCACTCCACTGACTCTG 3241
Db 3183 CAAGGAGAAACCTTTGAAGATGTGCTGGTCAAGATTTGACGGCACTCCACTGACTCTG 3242
Qy 3242 ATTTCCAAAGTTTCCCGCATGTTCTAGTCCGCTCTGACTAGACACACAAAGAGCTTCAAA 3301

Db 3243 ATTTCCAAAGTCTTCCCGCATGTTCTTAGTCGCCCTGACTAGACACACAAAGAGCTTCAAA 3302
Qy 3302 TATTACACCGTAGTGTAGATCCCTTTAGTACAGATAAATAGTAGTATGCTCTCTTTAAAGC 3361
Db 3303 TATTACACCGTAGTGTAGATCCCTTTAGTACAGATAAATAGTAGTATGCTCTCTTTAAAGC 3362
Qy 3362 TCCCTTCCCTTTAGAAATCTATATGTAGACAGCAGTAGTAGATACAGATTCAGATGAT 3421
Db 3363 TCCCTTCCCTTTAGAAATCTATATGTAGACAGCAGTAGTAGATACAGATTCAGATGAT 3422
Qy 3422 GCAGTGTTCAAAGGTCATAAATCTCTTTGTGGCAACACCTAAATCAGGAGACTTTCCAGAT 3481
Db 3423 GCAGTGTTCAAAGGTCATAAATCTCTTTGTGGCAACACCTAAATCAGGAGACTTTCCAGAT 3482
Qy 3482 CTACAGTCTTATACAGATGATAGCTCCCTGGTAAATAGTACTATCTATTAACAAGTAGAT 3541
Db 3483 CTGAGTCTTATATAGTATGATGCTCCCTGGTAAACAGTACTATATCTTAAACAAGTAGAT 3542
Qy 3542 GCTGTACCATGAGGTTACGGGTAATAGTCTTAATGTGAAGGATTTGTTCTTGAATTTT 3601
Db 3543 GCTGTACCATGAGGTTACGGGTAATAGTCTTAATGTGAAGGATTTGTTCTTGAATTTT 3602
Qy 3602 TCCAAAAGTATTCGATGCCAAAGAGGTGAAACCATGCTAGAGCCAGTTTTCGCTACC 3661
Db 3603 TCCAAAAGCATTCGATGCCAAAGAGGTGAAACCATGCTAGAGCCAGTTTTCGCTACC 3662
Qy 3662 GCGCGGAACCGCAAGGGCTGCAAGGCTACTCGAAATCTGGTTCGAATGATTAAGA 3721
Db 3663 GCGCGGAGAACCGCAAGGGTTCAGGACTACTCGAAATCTGGTTCGAATGATTAAGA 3722
Qy 3722 AATTTCACGACACAGACCTGACCGGGACGATTCACATTTGAGAGCACCGCATCTGTTGTA 3781
Db 3723 AATTTCACGACACAGACCTGACCGGGACGATTCACATTTGAGAGCACCGCATCTGTTGTA 3782
Qy 3782 GTAGATAAGTTTTTGTAGTACTATTTTATTAAGAAAGAAATACACAAAAATATTGCT 3841
Db 3783 GTAGATAAGTTTTTGTAGTACTATTTTATTAAGAAAGAAATACACAAAAATATTGCT 3842
Qy 3842 GAGGTGATGACGAAGGATTCATGATGAGATGGTT - GGAAGAACAGGAAGACTATT 3900
Db 3843 GAGGTGATGACGAAGGATTCATGATGAGATGGTTGGGAAACAGGAGAGTACTATC 3902
Qy 3901 GGACGACTTGCTTAACATACTTTACAGATCTCGCGGCCATCGATCAGTACAGCACAT 3960
Db 3903 GGAC - AGTTGGCTTAACATACTTTGTAGATCTCGCGGCCATCGACCAAGTATAAGCACAT 3961
Qy 3961 GATCAAGGCTCAACCAAAACAGAAATTCGACCTTTCAATTCCAGAAATGAATACCTGCTCT 4020
Db 3962 GATCAAGGCTCAACCAAAACAGAAATTCGACCTTTCAATTCCAGAAATGAATACCTGCTCT 4021
Qy 4021 GCAAAACAAATGCTTACCAATTCGAAGCAGATCAACCGGTATTTTGG - CCGGTTTTCTCAGA 4077
Db 4022 GCAAAACAAATGCTTACCAATTCGAAGCAGATCAACCGGTATTTTGGCCCGGTTTTTTCAGA 4081
Qy 4078 GCTTACAGGTTGCTGCTCGAGGCAATTTGATCTTAAGAAAGTTCTTTTCTTACTAGGAA 4137
Db 4082 GCTTACAGGTTGCTGCTCGAGGCAATTTGATCTTAAGAAAGTTCTTTTCTTACTAGGAA 4141
Qy 4138 AACTCCAGAACAGATTCAGAAATTTTCTCGGATCTCGACTCGACCTGCTCTATGAGATGT 4197
Db 4142 AACTCCAGAACAGATTCAGAAATTTTCTCGGATCTCGACTCGACCTGCTCTATGAGAGT 4201
Qy 4198 GTTAGAATCTGGATATTTCTAAGTATGATAAGTCA CAGAACGAGTTTCATTGTGCTGTA 4257
Db 4202 GTTAGAATCTGGATATTTCTAAGTATGATAAGTCA CAGAACGAGTTTCATTGTGCTGTA 4261
Qy 4258 GTATGAATATGGAAGATTTGGTCTCAATGAGTTTTTGGCCGAGTGTGGAAACAAGG 4317
Db 4262 GTATGAATATGGAAGATTTGGTCTCAATGAGTTTTTGGCCGAGTGTGGAAACAAGG 4321
Qy 4318 GCACAGAAACAACTTTGAAGGATTAATATGCTGGAAATCAAGACATGCTGTGCTGATCA 4377
Db 4322 ACACAGGAAACAACTTTGAAGGATTAATATGCTGGAAATCAAGACATGCTGTGCTGATCA 4381

Qy 4378 AAGGAAAGCGGTGATGTGACTACTTTTCATCGGCAATACTGTTATAATAGCAGCTTGCTT 4437
Db 4382 AAGGAAAGCGGTGATGTGACTACTTTTCATCGGCAATACTGTTATAATAGCAGCTTGCTT 4441
Qy 4438 GGGTTCAATGTTACCGATGGAAAGGTCATAAAAGTGCTTTTGTGGAGACGATTCGCT 4497
Db 4442 GGGCTCAATGTTACCTATGGAAAAGGTCATAAAAGTGCCCTTTTGTGGAGATGACTCCGT 4501
Qy 4498 TTTCTATTTTCCAAAGGTTTGGATTTCCCTGACATTCAGTCATGTGCTAATCTCATGTG 4557
Db 4502 TTTCTATTTCCGGAAGGTTTGGATTTCCCTGACATTCAGTCATGTGCTAATCTCATGTG 4561
Qy 4558 GAATTTTGAAGCCAAA CTGTATAGAAAAGAGTACGGTTACTTTTGTGGTAGATACATCAT 4617
Db 4562 GAATTTTGAAGCCAAA CTGTACAGAAAAGAGTACGGTTACTTTTGTGGTAGATACATCAT 4621
Qy 4618 ACACCATGATAAGGAGCAATAGTATATGATCCCTTGAAGTTGATCTCCAAACTGG 4677
Db 4622 ACACCATGATAAGGAGCAATAGTATATGATCCCTTGAAGTTGATCTCCAAACTGG 4681
Qy 4678 GGCAAAACATATCAAGGATTTATGATCACTTAGAAGAGTTAAGGCTGCTTTTGTGCGATGT 4737
Db 4682 GGCAAAACATATCAAGGATTTATGATCACTTAGAAGAGTTAAGGCTGCTTTTGTGCGATGT 4741
Qy 4738 TGTCTGTTCGCTCGAAA CTGGTCTTAGGCTTTCCGACGCTGAACCGAGCTATCAAGGA 4797
Db 4742 TGTCTGTTCGCTCGAAA CTGGTCTTAGGCTTTCCGACGCTGAACCGAGCTATCAAGGA 4798
Qy 4798 GGTTCATAAAACCGCGATTCGATGCTGCTTTTAAATGTTGTTAAACAAATTTTGTG 4857
Db 4799 GGTTCATAAAACCGCGATTCGATGCTGCTTTTAAATGTTGTTAAACAAATTTTGTG 4858
Qy 4858 TGATAAATTTTATTTAGAACTTTGTTTTAAATGGCTGTTAGTCTCAGAGACTCTGCA 4917
Db 4859 TGATAAATTTTATTTAGAACTTTGTTTTAAATGGCTGTTAGTCTCAGAGACTCTGCA 4918
Qy 4918 AAATTAGCGAGTTCATTTGATCTTTCGAAA CAGGATGAGTACTTCCGGCAATTCATGACTA 4977
Db 4919 AAATTAGCGAGTTCATTTGATCTTTTCGAAA CAGGATGAGTACTTCCGGCAATTCATGACTA 4978
Qy 4978 AGGTCAAGAGTGTAGAAATCGACTGTGACAGAGTATAGGCTGTTAAAGTACATAGTC 5037
Db 4979 AAGTCAAGAGTGTAGAAATCGACTGTGACAGAGTATAGGCTGTTAAAGTACATAGTC 5038
Qy 5038 TTTCTGATGTAGATTTTACTTTAAAGGTTTAAAGTTAGTTAAGAAAGGCTATGTGCTTAG 5097
Db 5039 TTTCTGATGTAGACTTACTTTAAAGGTTTAAAGTTAGTTAAGAAATGGGTATGTGCTTAG 5098
Qy 5098 CTGATTTGTTAGTGTCTGGGAGTGGAAATCTCCCGGATACTGCGGTGGTGTCAAGTG 5157
Db 5099 CTGATTTGTTAGTGTCTGGGAGTGGAAATCTCCCGGATACTGCGGTGGTGTCAAGTG 5158
Qy 5158 TTTGTTATTTAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAG 5217
Db 5159 TTTGTTATTTAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAGAGTAAAG 5218
Qy 5218 CCGCTGCTTCCAAAAGAAATTTTCTTTAAAGCTTAATCCCTAATTTTCAATAACATCCG 5277
Db 5219 CCGCTGCTTCCAAAAGAAATTTTCTTTAAAGCTTAATCCCTAATTTTCAATAACATCCG 5278
Qy 5278 AGGATGCTGAGAACGACCCGCTGGCAATGTGTAGTAAATATCAAGAGAGTGGCTATGGAAG 5337
Db 5279 AGGATGCTGAGAACGACCCGCTGGCAGGTGTAGTGAATATCAAGAGAGTGGCTATGGAAG 5338
Qy 5338 AAGGATGCTGCTTTTATCTTTTCGAGTTCGTTTTCAATTTTGTGTAGTACATAAAATTAATG 5397
Db 5339 AAGGATGCTGCTTTTATCTTTTCGAGTTCGTTTTCAATTTTGTGTAGTACATAAAATTAATG 5398
Qy 5398 TAAGAAAAGGTTTGAAGGAACTGATTTTGTAGTGTGACAGACGCTCGCCAAATTAAGACTCA 5457
Db 5399 TAAGAAAAGGTTTGAAGGAACTGATTTTGTAGAGTAAACAGACGCTTGCCTGATTAAGACTCA 5458

QY 5458 CTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACCAATGGCTGTGAAACTCGAAAAGG 5517
Db CTGAAAAGGTTGTTGAGGAGTTCGTGGATGAAGTACCAATGGCTGTGAAACTCGAAAAGG 5518
QY 5518 TTCGGGAAAACAAAAGAAATGTTAGTAAATTAATGTTAAATAAGAAAATAATAACA 5577
Db 5519 TCCGGAAAACAAAAGAGAGTGGTAGTAAATGTTAAATAAGAAAATAATAATA 5578
QY 5578 GTGGTAGAGAGGTTTAAATTTAGGNAATGAGGATAATGTAAGTATGACAGTCTA 5637
Db 5579 GTGGTAGAGAGGTTTAAATTTAGGNAATGAGGATAATGTAAGTATGACAGTCTA 5638
QY 5638 TCGGCTCATCAGTACGTTTAAATCAATATCCCTTATACAACTCAACTCTCCGAGCCAAAT 5697
Db 5639 TCGGCTCATCAGTACGTTTAAATCAATATCCCTTATACAACTCAACTCTCCGAGCCAAAT 5698
QY 5698 TGTTTACTTATCTTCCGCTTACGAGATCTCTGTCAGCTGATCAATCTGTGTACAAATGC 5757
Db 5699 TGTTTACTTATCTTCCGCTTATGAGATCTCTGTCAGCTGATCAATCTGTGTACGAATGC 5758
QY 5758 ATTGGGTAAACAGTTTCAAAGGAAACAGCTAGGACCAACAGTCCAAAGCAATTTGGCGGA 5817
Db 5759 ATTGGGTAAACAGTTTCAAAGGAAACAGCTAGGACCAACAGTCCAAAGCAATTTGGCGGA 5818
QY 5818 TGCCTCGAAACCTGTGCTAGTATGACAGTACGAGATTTCCCTGCATCGGATTTCTATGTGA 5877
Db 5819 TGCCTCGAAACCTGTGCTAGTATGACAGTACGAGATTTCCCTGCATCGGATTTCTATGTGA 5878
QY 5878 TAGATATAATTCGAGCTTGTATCGCTGATCAACCGCGGTTTAAATAGCTTTGATACCTAG 5937
Db 5879 TAGATATAATTCGAGCTTGTATCGCTGATCAACCGCGGTTTAAATAGCTTTGATACCTAG 5938
QY 5938 AAATAGATAATAGAGTTGATATCAACCGGACCGAATACTACTGAAATCGTTAAACGC 5997
Db 5939 AAATAGATAATAGAGTTGATATCAACCGGACCGAATACTACTGAAATCGTTAAACGC 5998
QY 5998 GACTCAGAGGTAGACGATCTAGTATGATCAAGGCTTCAATCAATTAATTTGGCTAA 6057
Db 5999 GACTCAGAGGTAGACGATCTAGTATGATCAAGGCTTCAATCAATTAATTTGGCTAA 6058
QY 6058 TGAATCGTTTCTGGAACCTGCGATGTTCAATCAAGCAGGCTTTGAGACTGCTAGTGACT 6117
Db 6059 TGAATCGTTTCTGGAACCTGCGATGTTCAATCAAGCAGGCTTTGAGACTGCTAGTGACT 6118
QY 6118 TGCTCGACCAAACTCCGGCTACTTACTATGTTGTGAGATTTCTTAAATAAAGTCG 6177
Db 6119 TGTTTGCACCAAACTCCAGCTACTTACTATGTTGTGAGATTTCTTAAATAAAGTCG 6178
QY 6178 CTGAGACTTAAATTCAGGCTGCTGATACCAAAATCAGCAGTGGTGTTCGTCACCTT 6237
Db 6179 CTGAGACTTAAATTCAGGCTGCTGATACCAAAATCAGCAGTGGTGTTCGTCACCTT 6238
QY 6238 AAATATACAGTTTGTATATCTGATCAACAGTTTAAACAGTGTGATGATGATG 6297
Db 6239 AAATATACAGTTTGTATATCTGATCAACAGTTTAAACAGTGTGATGATGATG 6298
QY 6298 GTATGGGTTAAACATCGAGAGTTCGAAATCTTCCCTTAACCCCGGCTAGCGGCCCA 6355
Db 6299 GTATGGGTTAAACATCGAGAGTTCGAAATCTTCCCTTAACCCCGGCTAGCGGCCCA 6356

RESULT 5

E07987
LOCUS E07987 6384 bp RNA linear PAT 29-SEP-1997
DEFINITION Tobacco mosaic virus L cDNA.
ACCESSION E07987
VERSION E07987.1 GI:2176118
KEYWORDS JP 1994233688-A/1.
SOURCE unidentified
ORGANISM unclassified
REFERENCE 1 (bases 1 to 6384)
AUTHORS Murai,F., Akiyama,T., Watanabe,Y. and Okada,Y.

TITLE
JOURNAL

TRANSCRIPTION VECTOR FOR IMPROVING PLANT VIRUS
Patent: JP 1994233688-A 1 23-AUG-1994;
SANKYO CO LTD
COMMENT
OS (Tobacco mosaic virus L)
PN JP 1994233688-A/1
PD 23-AUG-1994
PF 29-NOV-1993 JP 1993297704
PR 01-DEC-1992 JP 92P 322100
PI MURAI FUKASHI, AKIYAMA TOSHIYUKI, WATANABE YUICHIRO, PI
OKADA YOSHIMI
PC C12N15/83;
CC strandedness: Double;
CC topology: Linear;
CC Feature is identified by experimental;
FH Key Location/Qualifiers

FT source 1..6384
FT FT /organism='Unclassified'
FT 5'UTR 1..71
FT CDS 72..3422
FT /product='130kd-replication enzyme' FT CDS
FT 72..4919
FT /product=(seq:'tag', aa:Tyr)
FT CDS 4906..5700
FT /product='movement protein'
FT CDS 5703..6182
FT /product='coat protein'
FT 3'UTR 6183..6384.
FEATURES
source
1..6384
/organism="unidentified"
/mol_type="genomic RNA"
/db_xref="taxon:32644"

ORIGIN

Query Match 41.4%; Score 2631; DB 6; Length 6384;
Best Local Similarity 64.2%; Pred. No. 0;
Matches 4093; Conservative 0; Mismatches 2230; Indels 54; Gaps 7;
QY 18 CGACAAACAAATTAACAAACAAACAAACATATTACAAACAAACAAACAAACATGGCAC 77
Db 23 CAACAAACAAACAAACAAACAAACATTTACATTTTACATTC---TACAACTACATGGCAT 78
QY 78 ACATACATCTATAATTAAGCAACGCCCTTCTTGAAGCGTGAAGTGGTAAACAACTCTCG 137
Db 79 ACACAAACAGCACATCGTCGCTTGTCTTGAGACCGTCCGAGGTAAACATACCTTGG 138
QY 138 TTAATGACCTTGAAGAGCGGCATGTACATAGCGCGTGGGAAGATTTAAACCCCGCG 197
Db 139 TCAACGATCTTGCANAAGCGCGCTATATGACACAGCGGTCGATGAATTTAATGCTAGG 198
QY 198 ACCGTAGACCAAGGTTCAACTTTTCCAAAACTATTAGCGAAGAGCAACGCTTCTAGTCT 257
Db 199 ACCGAGGCTTAAAGTCAATTTTCCAAAGTAGTAAGCGAAGAACAGACGCTTATTCGAA 258
QY 258 CCAACGCGTACCGGAGTTCAGATTAACCTTTTATAATCTCAAAATCCGCTACACAGTT 317
Db 259 CCAAGCTTACCCAGATTTCCAAATTTACATTTTACACACGACAGATGCTGTGCAATCCC 318
QY 318 TGGCTGAGGTTTGAAGCAATTAAGATTTGAATATCTGATGCTACAAGTTCCCTTATGGAT 377
Db 319 TTGAGCGGCTCTCCGATCAATTAAGATTTGAATATCTGATGATGCAATTTCCCTTACGGAT 378
QY 378 CGCGCATATGATATAGTGGGAATTTTCAGACCAATTTGTTCAAGAGCGAGGATTAACG 437
Db 379 CATGACATATGATATCGGAGGTAATTTTGCATCTCATCTGTTCAAAGGGGAGCATACG 438
QY 438 TGCATTTGCTGATGCCCAATCTCGACATACGAGATATAATGAGCGACGAAAGGACAAAGG 497
Db 439 TTCACTGCTGATGCCGATCTGATGCTCCGACATATGCGGACGAGGAGGAGGAGG 498
QY 498 ACTCAATTGAGATGATATTTGTCAGGATTTGCTCTTAAACAGGTAATTTCTGAGTTTC 557

Db 499 ACAGTATTTGAACATATACCTTTCTAGGCTCGAGAGGGGCAACAAATCTCCAACTTCC 558
Qy 558 AAAGGGAGGCTTTTAAACAGGTATGAGAGCTCCCAACGAAGTCTGCTGCTCTTAAACTT 617
Db 559 AAAAGGAGGCTTTGACAGATACGCTGAATGCGCAACGAAGTAGTCTGTCAACGATACTT 618
Qy 618 TTCAGGATGTCGAATACATCGCCGACAGAGATAGTGTGAGAGATACGCTGTTGCTGCG 677
Db 619 TCCAAACGCTGAGGCATTTCTCAAGAAATGTTACACGGGAAGAGTGTATGCTATTGCTTGC 678
Qy 678 ACAGTTTCTATGATATCTCTGCGATGAGTTTGAGAGCTGCGTTAATATCTAAGAAATATAC 737
Db 679 ATAGTATATAGATATACCTCGCGACGAGTTTCGGCGGCGACTGCTGAGAAAGAAATGTAC 738
Qy 738 ATGATATGTTATGCACTTTCCTATTTTGGCAGAGCAATTAATCTAGACAGACGGAGTTA 797
Db 739 ATGATATGTTATGCGCTTTTCCACTTTTCCGAGAAATTTACTTCTCGAAGATTCACAGTCA 798
Qy 798 CGCTTAATGAATAGGCGCACTTTCAAAAGAGAGGTGATGTTCTTTTCTTTTCTTTG 857
Db 799 ACCTCGACGAGATCAATGCAATGTTTCCAAAGAGATGGAGACAGGTGTGACTTTTCTTTG 858
Qy 858 CTGATGAAGTACTTTTAAATTTATAGTCAATAATACAAAATATCTTGCAATATGTAGTTA 917
Db 859 CATCTGAGAGTACTTTTAAATTTATAGTCAATGTTATCTTAATATTTCTTAAGTATGTTGCA 918
Qy 918 AATCTTACTTCTGCTTCTAGTAGAATPAGTTTACTTTTAAAGGAATTTTACTCTACAGG 977
Db 919 AAACCTTACTTCCAGCCTCTAATAGAGAGTTTACATGAAGGAGTTTATTAGTAACATAG 978
Qy 978 TTAATACTGTTTGTAAATTTTAAATTTACCAAGTAGATACCTATATCTGTACAAAGTGTTA 1037
Db 979 TTAATACTGTTTGTAAATTTTCTAGATAGATACCTTTCTTAAATGTTTACAAAGGTGTAG 1038
Qy 1038 GACAGTAGGCTGTATGATGATCAGTTCTATGAGGCGATGGAAGCGCTTTTCTGTTACA 1097
Db 1039 CGCATAGGCTGTATAGTAGTACGCTTTTACAGGCTATGGAAGCGCATGGCACTACA 1098
Qy 1098 AGAAACCTTGGCCATGTTTCAACACTGAAAGAGCAATCTTTAGAGACAGCGCTTTGTTA 1157
Db 1099 AAAGACTCTTGGCATGTGCAACAGTGAAGAAATCTTGTAGAGGATTTCTTCAACAGTTA 1158
Qy 1158 ACTTTGGTCTTCAAGTAGAGCATGTTGATAGTACCGCTGTTTGGAGGTTCTATT 1217
Db 1159 ATTACTGTTTCCAAAATAGGGATATGGTATGTTCCACTATTTGACATATCTCTCG 1218
Qy 1218 CCAGCAAAAGATGACAAAGAGTGAAGTCAATTGTTAATCGTGACTTCGTTTACACAGTGC 1277
Db 1219 AGACTAGTAAAGAAACACGCAAGAGTCTTAGTTTCAAGGACTTTGTTTATACAGTGT 1278
Qy 1278 TTAATCATATCAGAACATATCAAGCCAAAGCGTTTAACTTACCAAGAGTATTTATCTTTCG 1337
Db 1279 TAAATCATATCGTACGTACCAAGGCGTAAAGCGCTTACTTACTCCAAAGTGTATCTTTTCG 1338
Qy 1338 TGGAGTCTTAAGATCCCGCTGATTAATCAATGTTGTTACTGCTAGTCTGAAATGGGATG 1397
Db 1339 TCGAATCAATTCGTTTCGAGAGTATCAATTAACGGGGTTACTGCTAGTCTGAGTGGGATG 1398
Qy 1398 TAGATAAAGCAATTTCTCAACCTTGTCAATGACTTTTCTTCTGCAAGTAAAGTGGCTG 1457
Db 1399 TCGATAATCATTTATACAGTCTTGTGTCGATGACGTTCTTCTTCTTCTTCTTCTTCTTTCG 1458
Qy 1458 CGCTTCAAGACGATATAGTAATGGGAAAGTTTCGGTGTCTGGAATGAAGCACTTCTGAAAC 1517
Db 1459 TTCTGAAAGACGATCTTTTGTATAGCAAGTTTGCATTTGGACCAAAAATCTGCTCTCAAC 1518
Qy 1518 TTATTTGGATGAGGTGGGCAATTTTGGAAACGTTTTCGCCACTATCAAGAGAGAT 1577
Db 1519 ATGTGCGGATGAGATTTCCCTAGCTTTCCGCAATGCTTTCCCATTCGATCAAGGAAAGAT 1578
Qy 1578 TGGTGACGAGGAAATTTCTGATGTAAGTGAGAAATGCTCTCAAGATCAAGATCCAGATC 1637

Db 1579 TGATAAACCCGAAACTGATCAAATATTACGGAGAAATGCGTTAGAGATCAGGGTGCCCGATC 1638
Qy 1638 TGATATGTCATGAGAAAGACAGGTTCTGTAGCTGAATACACCAAGTCTGAGAGATTACCGC 1697
Db 1639 TTTATGTCACTTTCCATGATAGGTTAGTTTCTGAGTACAAAATGTCAGTGACATGCGG 1698
Qy 1698 ATCTAGATATCAGAAGGCTTTAGAAGACTGAGCAAAATGCTAGACGCGTTATCAGAAAT 1757
Db 1699 TGTAGACATATTAGGAAAAAGATGGAAGAAATGAGGAAATGTACAATGCACTGTGCCGAAC 1758
Qy 1758 TATCTATCTTTAAGGGTGTCTGAATTTTCGATATCGCGAAAGTTCAAAGACATGTGCAAGG 1817
Db 1759 TGTCTGTACTTAAAAATTCAGACAAAGTTTCGATGTGACGTTTTCAGAGATGTCGAAT 1818
Qy 1818 CTTTAGATGTTAGTCTGTATGTCGACGACGAGTAATCGTGTGAGTGGCGGAGAAATAGAA 1877
Db 1819 CTTTAGAAGTTCGATCCAAATGACTGACGACAAAGGTAATAGTAGCAGTTTATGAGCAACGAGA 1878
Qy 1878 GCGGTTTAACTCTTACCTTTTGAATAAGCCAAACGAGGAGAAATGTGGCTAAAGCTCTTAAA - 1936
Db 1879 GTGCTCTTACTCTCACGTTTGAACAGCCCAAGCTAATGTTGCGGTAGCATGTGCAAG 1938
Qy 1937 -----AGCACGGCTCTGAGCCGCTGTGTATGCTTTGAACCGACATCCGAGAGAGTGAACG 1991
Db 1939 ATTCTGAAAGGCTTCTGATGGGCGTTGCTAGTTTACCTCAAGAGATGTTGAGGAACCGT 1998
Qy 1992 TAAATAAATTTTCTATTTCTGAGAAAGGAGATGTCCTGTGTGTGTCGAGAAAGTCAATGTT 2051
Db 1999 CCATAAAGGTTTCGATGGCCGCTGCTGATGTTTACAAATTTGGCGGATTAATCTGGCGACGTT 2058
Qy 2052 TGACGAATGCTTAATAGACACAGGAGTTGGAGTCCCTCAACGATTTCCATTAAGGCTT 2111
Db 2059 CTGAATCTTCAATACATAGGAGCGAGAGATGAGTCTCTCGAGAGTTTTCATATGGCAA 2118
Qy 2112 GCGTGAAGTGTGATTAACAAAGCAATGGCATCGGTTGTCTACACTGGCTCACTCAAG 2171
Db 2119 CAGCTAGTCTGTTAAATTCATAAGCAGATGTTTCGATCGTGTACACGCGGCTCTTAAAG 2178
Qy 2172 TTCAAAATTAAGAACTATGTGGAAGTTTGGCAGCTTCTGTTGTCGCGCACTGTATCAA 2231
Db 2179 TTCAAAATTAAGAACTATGTATAGACAGCTGCTGCTGCTCTCTGCTGCGGTGCGA 2238
Qy 2232 ATCTATCAAGTCACTAAAGATGAAGTGAAGTATGATTTCTGATTTCCAGGAGAGAAAGTTG 2291
Db 2239 ATCTAGTGAAGATCTTAAAGATACACCGCGATGAGCTTGAACCTCGTCAAAAGTTG 2298
Qy 2292 GTGTTTGGGATGTCACTTTGAAAAGTGGCTCTCAAAACCTCGCGCCAAAGGTCATTCAT 2351
Db 2299 GAGTTCTGGATGTTGCTTCGAAAAGGTTGGCTAGTTAAACCATCCGCAAGAACCATGCAT 2358
Qy 2352 GGGGAGTTGCTGGATTAACAGGGGAAATGTTTACTGCACTTCTATCTTATGCAAGGAG 2411
Db 2359 GGGGGTGTGTGAGACTCATCGAGGAAATATACGTCGCAATTAATCTGGAGCACGATGAAT 2418
Qy 2412 ATAGAATGTTGACTGAGAGGACTGAGAGGAGTGGCTGTATCATCTGATACAAATGTTAT 2471
Db 2419 TTGCAATTTATCTGCGATTAATCGGCAAGGTTGGCTGTGAGTTCTGAGTCTGAGTAGTAT 2478
Qy 2472 ATTCTGATATTGCAAAAGCTCCAAAATCTGAGGAAACAAATGAGAGACGGTGAACCCACG 2531
Db 2479 ATTCTGATATGGCTTAAACTCAGGACTCTGAGAGATGTTGCTCAAAAGATGGAACCAACG 2538
Qy 2532 AACCTACTGCAAGAGATGTTACTTGTGATGGGTGCTGTTGTGTTGGAAGTATGACAAAGGAG 2591
Db 2539 TTAGTTTCAGCAAGGTTGTTTGTGTGATGCGTTCAGGGTTCGGGAGCAAAAGGAAA 2598
Qy 2592 ATTTTGAAGATTTGATCTTGTAGGAGATTTGATCTTGTGTTCTCTGGAATAACAAGCTGCTG 2651
Db 2599 TTCTTTGAGAGTTTAAATTTTCGAGAGATCTAATTTCTTGTCCCTGCTGCTCAAGCTGCG 2658
Qy 2652 CTATGATCAGAAGAGGCTTAAATTCATCTGAGCTGATGAGGCAACCAATGAGCAATGTGA 2711
Db 2659 AGATGATCAGAAGAGAGCTAATGCTCGGCGCAATATAGTGGCTTACAAAGGATTAATGTGC 2718

Db 4876 ATAAAGTTTGTGTTAGAACTTATTTCTTGTAGTGGCTCTAGTTGTTAAAGGTAAGGTAAT 4935
QY 4920 ATTACGAGTTCATGATCTTTCGAAACAGAGTATGATCTTCGGCATTTCATGACTAAG 4979
Db 4936 ATTAATGAGTTTATCGATCTGTCNAAGTCTGAGNAACCTTCTCCCGTCGATGTTTCAGGCT 4995
QY 4980 GTCAAGAGTGTAGAAATTCGACTGTGGAAGAATTTATGGCTGTTAAAGATGATAGTCTT 5039
Db 4996 GTAAAGAGTGTATGTTTCAAGGTTGATAGATTTAGTCCATGAAATGAATCATTTG 5055
QY 5040 TCTGATGTAGATTACTTAAAGGTTTAAAGTTAGTTTAAAGAGGTAATGTTGCTTAGCT 5099
Db 5056 TCTGAAGTAAATCTCTTAAAGGTTTAAACCTTATAGAGGTTGGGTATGTTTGTCTTAGTC 5115
QY 5100 GATTTGGTATGCTCTGGGAGTGGAACTCTCCGGATATCTCCGTTGCTGGTGTCTAGTCTT 5159
Db 5116 GGTCTGTTGTCTCCGGTGAAGTAAATTTACAGATAATTTGCGGTGGTGTGATGTC 5175
QY 5160 TGTATTGTAGATAAGAGATGAAGAGGAGTAAGGAAGCAACGCTGGGTGCTATCAGCC 5219
Db 5176 TGCATGTTGACAAGAGATGAAGAGCGGACGACCACTGGGTCTATTTACACT 5235
QY 5220 CTTGCTTGCACAAAGAAATTTTCTTTAAAGCTAAATCCCTAAATTTTCAATTAACATCCGAG 5279
Db 5236 GCTGCTGCTAAAGCGGTTTCAGTTTAAAGTGGTCCCAATTTACGGTATTACAAACAAAG 5295
QY 5280 GATGCTGAGAAGACCCGTGGCAAGTGTAGTGAATATCAAGAGGAGTGGCTATGAGAA 5339
Db 5296 GATGAGAAAGAACATATGCGAGGCTCTTAGTAAATTTTAAATAATGTAATAATGATGCG 5355
QY 5340 GATAGTGTCTTATCTTTGAGTTCGTTTCAATTTGTTGATGATACATAAAATATGTA 5399
Db 5356 GGTCTGCTCCCTTGTCTATAGATTTGCTGCTGTTGTTGTTTATTAATAATATATA 5415
QY 5400 AGAAAGAGTTTGAAGGAACTATTTTGAAGTGTGACAGCGGCTCGCAATTTGAATCACT 5459
Db 5416 AATTTGGTTTGAAGGAAAGTAAAGAGTGTGAACGATGGAGGACCCATGGAATTTCCG 5475
QY 5460 GAAAGGTTTGAAGAGTTCGTTGATGATGATGATGATGATGATGATGATGATGATGAT 5519
Db 5476 GAAGAGTGTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5535
QY 5520 CCGGAAACAAAGAAATGTTAGTAAATGTTAAATTAAT 5561
Db 5536 CGAACCAATCTCAAAAGAGTCCGAAATTAATTAATTAATTAATTAATTAATTAATTAAT 5595
QY 5562 ---AAGAAATAAATACAGTGTGTAAGAGGGTTTTTAAATTTGAGGAAATTTGAGGATAAT 5618
Db 5596 GCGGAAGGCTAAACCAAAAGTTTGTATGAAAGTTTGAAGAGAGTTTGTATTAATTTGAT 5655
QY 5619 GTAAGTATGACAGTCTATCGGCTCATCGGTACGTTTATCAATATGCTTATACAA 5678
Db 5656 GAAGATGAAGCCGAGACGTCGTCGCGGATTCGATTTCGTTATTAATATGCTTACTCAA 5715
QY 5679 TCAACTCTCGACCAATTTGTTTACTTATCTTCGCTTACGAGATCTCTGTCAGCTGA 5738
Db 5716 TCACTTCTCATCGCAATTTGTTTGTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTTGTT 5775
QY 5739 TCAATCTGTGTACAAATTCGTTGGTAAACAGTTTCAAGCGCAACAGCTTAGGACAACAG 5798
Db 5776 TAAACGTTGTACAAATTCGTTAGGTAAACAGTTTCAACACAGCAAGCAAGCAACTACTG 5835
QY 5799 TCAACAGCAATTTGCGGATGCTGGAACCTGTGCTGTAGTATGACAGTGAATTTCTTG 5858
Db 5836 TTCAACAGCAGTTTCAGCGAGTGTGGAACCTTTCCCTCAGACACCGCTCAGATTTCTTG 5895
QY 5859 CATCGGATTTCTATGTATAGATAATAATTCGACGCTTGTATCGCTGATCAGCGGCTTAT 5918
Db 5896 GCGATGTTTATAAGGTGTACAGGTACATGCGAGTTTATGATTTCTTAATTTACTGCGTTG 5955
QY 5919 TAAATAGCTTTGATCTAGAAATAGATAATATAGAGGTTTGAATTAATCAACCCGACCGAATA 5978

Db 5956 TGGGGGCTTTGATCTAGGAATAGATAATTCGAAGTAGAATAACAGACAGAGTCCGACAA 6015
QY 5979 CTACTGAATCGTTAAACCGGACTCAGAGGGTAGACGATGCTACTGTAGCTAAGGGCTT 6038
Db 6016 CAGCTGAAACGTTAGATGCTACCCGCGGGTAGACGCTACGGTTGCAATTCGGTCTG 6075
QY 6039 CAATCAATAATTTGGCTAATGAACCTGGTTCGTTGGAACCTGGCATGTTTCAATCAAGAGGCT 6098
Db 6076 CTATAATAATTTAGTTAATGAACCTAGTAAGAGGTAAGTGGACTGTACATCAGAAATACTT 6135
QY 6099 TTGACAGCTAGTGGACTGCTGCGACCAACCTCCGCTACTTACTTAGCTTATTTGTTGTAG 6158
Db 6136 TTGAAGATATGCTGGGTTGGTCTGACCTCTGCACTTCATCTTAAATGCAATAGTGTCT 6195
QY 6159 ATTTCTTAAATAAAGTCTGCAAGACTTTAAATTTCAAGGTGGCTGATACCAAAATCAAC 6218
Db 6196 GAAATATAA-----GTTTGTGTTTCTTAAACACACGCTGTAGTACGATAACGTAC 6247
QY 6219 AGTGTGTTGCTGCTCACTTAAATAATGAATTTGCTATCTGATTCATCAACAGTTAAACCA 6278
Db 6248 AGTGTGTTTCCCTCCACTTAAATCGAAGGTTAGTGTCTTTGGAGCGCGGAGTAAACATA 6307
QY 6279 TGTGATGTTGATATCTGTTGATGCGTAAACATCGGAGAGGTTTCAATCTCCCTAA 6338
Db 6308 TATGTTTCATATATGTCGTAGGACGTTAAATAAAGCGAGGATTCGAATTCCTCCCGAA 6367
QY 6339 CCGCGGTAGCGGCCCA 6355
Db 6368 CCCCCGTTGGGCCCA 6384

RESULT 6

TOTMV8 6384 bp RNA linear VRL 10-FEB-1999
Tobacco mosaic virus tomato strain (L) genome.
LOCUS X02144
ACCESSION X02144.1 GI:62128
VERSION
KEYWORDS
SOURCE
ORGANISM
Tobacco mosaic virus
Tobacco mosaic virus
viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE 1 (bases 1 to 6384)
Okada, Y.
AUTHORS Ohno, T., Aoyagi, M., Yamanashi, Y., Saito, H., Ikawa, S., Meshi, T. and
Nucleotide sequence of the tobacco mosaic virus (tomato strain)
genome and comparison with the common strain genome
J. Biochem. 96 (6), 1915-1923 (1984)
MEDLINE 85157522
PUBMED 6549393
REFERENCE 2 (bases 1 to 6384)
Belenovich, E., Genetsov, E., Novikov, V. and Zavriv, S.
AUTHORS Properties and the genome structure of the K2 strain of tobacco
mosaic virus
Mol. Biol. (Mosk.)
COMMENT Data kindly reviewed (12-FEB-1986) by Y. Okada.
FEATURES
Location/Qualifiers
1. .6384
/organism="Tobacco mosaic virus"
/mol_type="genomic RNA"
/strain="OM strain."
/db_xref="taxon:12242"
72. .4922
/notes="unnamed protein product; readthrough stop codon
180 K protein"
/codon_start=1
/transl_except=(pos:3420. 3422,aa:OTHER)
/protein_id="CAA26085.1"
/db_xref="GI:2440217"
/db_xref="GOA:P03587"
/db_xref="Swiss-Prot:P03587"
/translation="MAYTQTATSSALLETVRGNTLVNDLAKRLRYDTAVDEFNARDR
RPKNFVKVSEQTLLIATKAPFPQITNTQNVHSLAGLSLELVLMQIPVG
SLTYDIGGNFASHLFGKRAYVHCMPNLDVDRIMRHEGQKDSIELYLSRLRGNKHPV

source

CDS

Db 979 TTAATACCTGCTTTGTAATTTCTAGATAGATACCTTTCTTATTTGACAAAGGTGTAG 1038
Qy 1038 GACAGTAGGGTGTGATAGTATCAGTTCTATAGGCGATGGAAGCCCTTTGCTTACA 1097
Db 1039 CGCATAGGGGTGTAGATAGTACAGGTTTACAGGCTATGGAAGCCATGGCACTACA 1098
Qy 1098 AGAAACCTTGGCCATGTTCAACACTGAAAGAGCAATCTTTAGAGACACGGCTTCGGTTA 1157
Db 1099 AABAGACTCTTGGATGTGCAACAGTGAAGAACTCTGTAGAGGATCTTTCATCAGTTA 1158
Qy 1158 ACTTTGGTCCCTTAAGTAGAAGACATGGTGTATAGTACCGCTGTTGAGGGTCTTATTA 1217
Db 1159 ATTAAGTCTTCCAAATAGGGATATGGTATAGTTCCATATTTGACATATCTCTCG 1218
Qy 1218 CCAGCAAAAGATGACAGAGGTAGGTCAATTTGTAATCGTGACTTCGTTTACACAGTGC 1277
Db 1219 AGACTAGTAAGAAACACGCAAGAGGTCTTAGTTTCAAGGACCTTGTGTTATACAGTGT 1278
Qy 1278 TTAATCATATCAGAACATATCAAGCCAAAGCGTTAACTTACAGAACGATATTAATCTTCG 1337
Db 1279 TAAATCATATCGTAGTACCAGGCCAAAGCGCTTACTTACTCCAAAGCTTATCTTTTCG 1338
Qy 1338 TGGAGTCTAATAGATCCCGGTGATPAATCAATGTTGTTACTGCTAGGTCTGNAATGGATG 1397
Db 1339 TCGAATCAATTCGTTTCGAGAGTGATCATTAACGGGTTTACTGCTAGGTCTGAGTGGGATG 1398
Qy 1398 TAGATAAGCAATCTTCAACCTTGTCAATGACTTCTTCTGTCAGACTAAGCTGGCTG 1457
Db 1399 TCGATAAATCATTAATACAGTCTTGTGATGACGCTTCTTCTCAACACCAAGCTTGGC 1458
Qy 1458 CGCTTCAAGACGATATAGTAATGGAAAGTTTCGGTGTGTTGGATAGACCACTTCTGAAC 1517
Db 1459 TTCTGAAGACGATCTTTTGATTAGCAAGTTTGCACTTGGACCAAACTGTCTCAAC 1518
Qy 1518 TTATTTGGATAGGTGGCAAAATTTTGGAAACGTTTTCCTCCACTATCAAGAGAGAT 1577
Db 1519 ATGTGTGGATGAGATTTCCCTAGCTTTCGGCAATGCTTTCCTCATCGATCAAGGAAAGAT 1578
Qy 1578 TGGTACGAGGAAATCTCGATGTAGTGAAGTCTCTGAAGATCAAGATCCAGATC 1637
Db 1579 TGATAAACCGGAAATCTGATCAAAATTCAGAGAAATCGGTTAGAGATCAGGTTGCCGATC 1638
Qy 1638 TGATATGTACATGGAAGACAGGTTCTGATGCTGAATATACCAAGTCTGAGGAGTTACCGC 1697
Db 1639 TTATGTCACTTCCATGATAGTTAGTTTCTGAGTACAAATGTCAAGTGGACATGCCGG 1698
Qy 1698 ATCTAGATATCAAGAGGCTTAGAAGAGCTGAGCAAAATGTACAGCGGTTATCAGAT 1757
Db 1699 TGCTAGACATTAGGAAAGATGGAAGAACTGAGGAAATGTACAAATGCACTGTCCGAAC 1758
Qy 1758 TATCTATCCTTAAGGTCGTGATTAATTCGATATCGGAAGTTCAAGACATGTGCAAGG 1817
Db 1759 TGCTGTACTTAAATTCAGACAAAGTTCGATGTGACGCTTTTCCAGATGTGCCAT 1818
Qy 1818 CTTTATAGTGTAGTCTGATGTGGACACAGATTAATCGTTCAGTGGCGGAGATAGAA 1877
Db 1819 CTTTGAAGTCGATCCATGACTGCGACAAAGGTAATAGTACGATTTAGCAACGAGA 1878
Qy 1878 GCGGTTTAACTCTTACTTTTGTATAGCCAAACGAGGAGAAATGTGGCTTAAGGCTCTTAAA- 1936
Db 1879 GTGGTCTTACTCTACGTTTGAACAGCCACCGCAAGCTAATGTGGCTAGCAATTTGCAAG 1938
Qy 1937 -----AGCACCGGCTGAGCGCGTGTATGTTGAAACCGACATCCGAAGAGGTGAACG 1991
Db 1939 ATTCGTGAAGGCTTCTGATGGGCGTGTGGTAGTTACCTCAAGAGATGTGAGAACCGT 1998
Qy 1992 TAAATAAATTTCTATGTGTGAGAAAGGAGATTTGCCCTGTGTGTGAGAAAGTCTATGTT 2051
Db 1999 CATAAAGGTTTCGATGGCCCGTGTGTAGTTACATTTGGCCGATTTATCTGGCGAGCTTC 2058
Qy 2052 TGACGAATGCTAATCTAGAGCACACGAGGTGAGTCCCTCAACGATTTCCATAGGCTT 2111

Db 2059 CTGAATCTTCATACACTAGGAGCGAGGAGATTGAGTCTCTCGACAGTTTTCATATGGCAA 2118
Qy 2112 GCGTGGATAGTGTGATTAACAAAGCAATGGCATCGGTTGTCTACTGCTGGCTCACTCAAG 2171
Db 2119 CAGCTAGTTCTGTTAATTCATAAGCAGATGTCTTCGATCGTGTACACGCGCCTCTTAAAG 2178
Qy 2172 TTCAACAAATGAAGAACTATGTGACACGATTTGGGCGACTTCGTTGTCGCGCACTGTATCAA 2231
Db 2179 TTCAACAAATGAAGAACTTTATAGACACCTGGTAGCCTCGCTCTCTCTCGGGTGTGCA 2238
Qy 2232 ATCTATGCAAGTCTCAATAAGGATGAAGTCGGGTATGATTTCTGATTTCCAGGGAGAAAGTTG 2291
Db 2239 ATCTATGCAAGTCTCAATAAGATACAGCGCGGATTTGACCTTGAACCTGTCATAAAGTTTCG 2298
Qy 2292 GTGTTGGGATGTCACTTTGAAAAAGTGGCTCTCAAACTGGCGGCCAAAGGTCAATTCAT 2351
Db 2299 GAGTCTCGATGTGCTTCGAAAAAGGTGGCTAGTTAAACCATCCGCAAGAACCATGCTAT 2358
Qy 2352 GGGGAGTTGCTCGATTTACAGGGGAAATGTTTACTGCACTTCTATGTAAGAGAG 2411
Db 2359 GGGGGTTGTTGAGACTCATCGAGGAAATATACGCTCGCATTTACTGGAGCAGATGAT 2418
Qy 2412 ATAGAATGGTCACTGAGAGCGACTGGAGGAGGGTGGCTGTATCATCTCATACAATGGTAT 2471
Db 2419 TTGCAATTTATTCGTCGATTAATCTGGCGACGGTGGCTGTGAGTTCTGAGTCGGTAGTAT 2478
Qy 2472 ATTCGTATTTGCAAAAGCTCCAAATCTGAGGAAACAAATGAGAGACGGTGAACCCACG 2531
Db 2479 ATTCGTATATGGCTAAACTCAGGACTCTGAGAAAGATTGCTCAAGAGTGGAGAACACACG 2538
Qy 2532 AACCTACTGCAAGATGTTACTTGTGATGGGTGGCTGCTGTTGTGGAAGTACAAAGGAG 2591
Db 2539 TTAGTTACGAAAGGTGGTTTGGTGGATGGCTTCCAGGGTGGCGGAGACAAAGGAAA 2598
Qy 2592 ATTTTGAAGATTTCATCTGATGAGGATTTGATCTTGTCTTCTCGAAAAACAAGCTGCTG 2651
Db 2599 TTCTTTCAGAGTTAATTTTCGAGAGATCTAATTTCTTGTCCCTGGTCTCAAGCTGCCG 2658
Qy 2652 CTATGATCAGAGAAAGGCTAATTCATCTGATGATGAAGAGCCCAATGGAATGTGA 2711
Db 2659 AGATGATCAGAGAAAGAGCTAATGCTCGGCGATTAATAGTGGCTACAAAGGATAATGTGC 2718
Qy 2712 GAACTGATATTCATCTTAATGATCCAAACCGGATCAC-----ACAAGAGGC 2762
Db 2719 GCACGCTCGATTCATTTTGTGATTAATACGGGAAAGGCGACGCTGTCTAGTTCCAAAGAT 2778
Qy 2763 TTTTATTTGATGAAGGTTGATGTCACACCGGTTGTGTTAACTTCTCTGGTCTTATCT 2822
Db 2779 TGCTCATAGACGAGGTTTGTGCTGCATACGTTGTGTTGTGAAATTTTGGTTGAAATGT 2838
Qy 2823 CTGTTGCGACATCCGATACATTTACGGAGATACACAGAGATTCCTTTTCATTAACAGAG 2882
Db 2839 CTCTGCGGATATTCATATGTTTATGGAGACACCCAAACAGATTCGTTACATCAACAGAG 2898
Qy 2883 TTCAGAAATTTCCGTTATCCCAACATTTTGAAGCTGCAAGTGGATGAAGTTGAGATGA 2942
Db 2899 TAACTGGTTTCCGTTACCCCTGCACACTTTGCAAAATTTGAGGTCGACGAGTTCGAACAA 2958
Qy 2943 GGAGACCACTCTGAGATGCCCGAGGTGATGTCAAATTTTTCCTCAATCAAGATGACGAAG 3002
Db 2959 GAAAGACTACTCTCTGTTGTCGGCTGATGTACACACTTCTCTTAATCAAGATGATGAAG 3018
Qy 3003 GAGCGGTGACAAACCTTCAACTGTATCAACAGATTCGGTCTCATCTGAGATGATAGCGGTTA 3062
Db 3019 GACAGTAAATGTGACGCTCTCTGAAAAAGAAATCAGTTTCCAGGAAATGGTTAGTGGGG 3078
Qy 3063 AGGAGTACTAAACAGTGTTCCAACCACTAAAGGGGAAATTTGAACTTTTCACCTCAGG 3122
Db 3079 CTGCGTCTAATCTGTTCCAGCGCTTAAAGGGAAAAATTTTGACTTTTTCACACAGT 3138
Qy 3123 CTGATAAATTTGAGTTAGAGGAGAGGCGCTATAAGAAATGTGAACACCGTTTCATGAGATCC 3182
Db 3139 CTGACAGAGGAGGCGCTTCTCTCAAGGGGCTACCGAGATGTCCATATCTGTACATGAGGTAC 3198

Db 1580 TGATAAACCGGAAACTGATCAAAATTACGAGAAATTCGTTAGAGATCAGGGTGCCCGATC 1639
Qy 1638 TGTATGTCACATGGAAGACAGAGTTCTGTAGCTGTAATACACCAAGTCTCAGAGATTACCGC 1697
Db 1640 TTTATGTCACATTTCCATGATAGTTAGTTCTGAGTACAAATGTTCAGTGACATGCCCG 1699
Qy 1698 ATCTAGATATCAAGAAAGACTTAGAAGAGCTGAGCAAAATGTACGACCGTTATCAGAA 1757
Db 1700 TGCTAGACATTTAGGAAAAAGATGGAAGAACTGAGGAAATGTACAAATGCATGCTCCGAAC 1759
Qy 1758 TATCTATCTTAAGGGTCTGATTAATTTTCGATATCGCAAGTTCAAAGACATGTGCAAG 1817
Db 1760 TGTCTGTACTTTAAAAATTCAGACAAATTCGATTTGACGTTTTTTCCAGATGTGCCAAT 1819
Qy 1818 CTTTATAGTTTGTCTGATGTCGAGCAGCAGTAAATCGTTGCACTGCGCCGAGAAATAGAA 1877
Db 1820 CTTTGAAGTTCGATCCAATGATCTGAGCAAAAGGTAATATAGTACGATTTATGACAAAGAGA 1879
Qy 1878 GCGGTTTAACTCTTACTTTTATAGCAACCGAGGAGAAATGTGGCTTAAGGCTCTTAAA- 1936
Db 1880 GTGGTCTTACTCTCAGCTTTGAACAGCCCAAGAGCTAAATGTTGCGCTAGCATTTGCAAG 1939
Qy 1937 -----AGCAGCGGCTCTGAGCGGTGATATGTTTGAACCGACATCCGAAGAGTGAACG 1991
Db 1940 ATTCTGAAAAGGCTTCTGATGGGCGTTGATAGTTTACCTCAAGAGATGTTGAGGAAACCGT 1999
Qy 1992 TAAATAAAATTTCTATTCGTGAGAAAGGAGATTCGCTGTGTGTCGAGAAAGTCAATGTT 2051
Db 2000 CCAATAAAGGTTTCGATGCGCCGTGTGAGTTTACAATGCGCGGATTAATCTGGCGACGTT 2059
Qy 2052 TGACGAATGCTAACTTAGAGCAACCGAGGAGTTGGAGTCCCTCAACGATTTCCATTAAGGCTT 2111
Db 2060 CTGAATCTTCATACACTAGGAGCGAGGAGATTGAGTCTCTCGAGCAGTTTCATATGGCAA 2119
Qy 2112 GCGTGGATGTGATTAACAAGCAATGCGCATCGGTTGTCTACTGCTCCTCACTCAAG 2171
Db 2120 CAGCTAGTTCTGTTAATTCATAGCAGATGTGTTCGATCGTGTACACGCGGCCCTCTTAAAG 2179
Qy 2172 TTCAACAAATGAAGAACTATGAGCAGTTTGGCAGCTTCGTTGTCGCACTGTATCAA 2231
Db 2180 TTCAACAAATGAAGAACTTATAGACAGCTGTGAGCTCGCTCTCTGCTGGGTGTGCA 2239
Qy 2232 ATCTATGCAAGTCACTAAAGGATGAAGTCGGGTATGATTTCTGATTTCCAGGGGAAAGTTG 2291
Db 2240 ATCTAGTGAAGATCTTAAAGATACAGCTCGATTTGACCTTGAAACTCGTCAAAAGTTTCG 2299
Qy 2292 GTGTTTGGATGTCACTTTGAAAAGTGGCTCTCAAACTCGCGCCCAAGGTCAATTCAT 2351
Db 2300 GAGTTCTGGATGTTCTTCGAAAAGGTGGCTAGTTAAACCATTCGCAAGGACCATGCA 2359
Qy 2352 GGGGAGTTGTCCTGGATTAACAGGGGAAAATGTTTACTGCACTTCTATCTTATGAAGGAG 2411
Db 2360 GGGGGTTGTTGAGACTCATGCGAGGAATATCAGCTCGCATTTCTGGAGCAGATGAAT 2419
Qy 2412 ATAGAATGGTGACTGAGAGCGACTCGAGAGGGTGGCTGTATCATCTGATACAATGGTAT 2471
Db 2420 TTGGCATTATTACGTGCGATACTGCGACGGGTGGCTGTGAGTTCTGAGTCGGTAGTAT 2479
Qy 2472 ATTCTGATATTCGAAGTCCAAATCTGAGGAAACATGAGACCGGTGAACCCACG 2531
Db 2480 ATTCTGATATGCTAAACTCAGGACTCTGAGAAGATTTGCTCAAGATGGAGAACACACG 2539
Qy 2532 AACCTACTGCAAGATGTTACTTTGATGGGGTGGCTGTTGTGGAAGTACAAAGGAG 2591
Db 2540 TTAGTTACGAAAGGTGGTTTGGTGGATGGGTTCCAGGGTTCGGGAAGACAAAGGAA 2599
Qy 2592 ATTTTGAAGATTGATCTTGATGAGGATTTGATCTTGTTGTTCTTGAAAAAAGTGTG 2651
Db 2600 TTCTTTTCGAGAGTTAATTTGGAAGAAGATCTAATTTCTTGTCCCTGCTCAAGCTGCG 2659
Qy 2652 CTATGATCAGAGAGGCTAATTTCTGACTGATAGAGCCCAATGGCAATGTGA 2711
Db 2660 AGATGATCAGAGAAGAGCTAATGCGTCGGGCATAATAGTGGCTACAAAGGATAATGTGC 2719

Qy 2712 GAACGGTAGATTCACTTCTTAATGCATCCAAACCGCGATCAC-----ACAAGAGC 2762
Db 2720 GCACCGTTGATTCAATTTTGTGATTAATACGGGAAAAGGCGACGCTGTCAAGTTCAAAAGAT 2779
Qy 2763 TTTTATTTGATGAAGGGTTGATGCTGCACACCGGTTGTGTAACTTCTCGTGGTCTTACT 2822
Db 2780 TGTTCATAGACGAAGGTTTGTATGCTGCATCTGGTTGTGTGAATTTTGTGTGAAATGT 2839
Qy 2823 CTGTTGCGACATCGCATACATTTACGGAGATACACAGCAGATTCCTTTCAATTAACAGAG 2882
Db 2840 CTCTGTGCGATTTGCAATATGTTTATGGAGACACCCACAGATTTCCGTACATCAACAGAG 2899
Qy 2883 TTCAGAAATTTCCCGTATCCCAAAACATTTTGAAGAGCTGCAAGTGGATGAAGTGAATGA 2942
Db 2900 TAACTGGTTTCCCGTACCCTGCACACTTTTGCAAAATTTGGAGGTGCAGAAATCGAACA 2959
Qy 2943 GGAGACACACTGAGATGCCAGGTGATGAATTTTCTTCAATTCGAATGACGAG 3002
Db 2960 GAAGAACTACTCTTCTGTTGTCGGCTGATGTGCACACTTCTTAATCAAAAGGTATGAAG 3019
Qy 3003 GAGCGGTGACAAACCACTTCAACTGTACAAACGATCGGTCTCATCTGAGATGATAGCGGTA 3062
Db 3020 GACAGTAAATGTGACGCTCTTCTGAAAAGAAATTCAGTTTCCAGGAAATGTTAGTGGG 3079
Qy 3063 AGGAGTACTAAACAGATGTTTCCAAACCACTAAAGGGGAAAATTTGAATTTTCACTCAGG 3122
Db 3080 CTGCGTCTATCAATCTCTGTGTTCAAGCGCTTAAAGGAAAATTTGACTTTTCAACAGT 3139
Qy 3123 CTGATAAATTTGAGTAGAGGAGGAGGCTATAGAAATGTGAACACCGTTTCATGAGATCC 3182
Db 3140 CTGAAGAGGAGGCGCTTCTCTCAAGGGCTACGCAAGATGTCCATCTGTACATAGGTAC 3199
Qy 3183 AAGGAGAACTTTGAAGATGTGCTGCTGCTGAGATTGACGCAACTCCACTGACTGCA 3242
Db 3200 AAGTGAGACTATGACAGCGTATCGTTAGTTTCGATTAACCTACGCTGTATCTATCA 3259
Qy 3243 TTTCCAGTCTTCCCAGCATGTTCTAGTCGCTGACTGACTAGACACACAAGAGCTTCAAT 3302
Db 3260 TCGAAGAGAGACAGTCCGCAATGTTCTGCTCTGTTCAAGACACACAATAATCCCTAAAGT 3319
Qy 3303 ATTACACCGTAGTGTAGATCTTTAGTACAGATTAATAGTATGTTGTTCTTTTAAAGCT 3362
Db 3320 ACTACCGTTGTGTAGTATCTTTAGTTAGTATCATTTAGAGATTTAGAACGGTTAGTA 3379
Qy 3363 CTTCTTTTGAATAATGATATGTTAGAGCAGGTAGTAGAGCAATTTACAGATGGATG 3422
Db 3380 GTTACTTATTAGACATGTACAAAGTAGATGAGGTACTCAATAGCAATTCAGGTCGACT 3439
Qy 3423 CAGTGTCAAAGGTCAATAATCTCTTTTGGGCAACACCTTAATTCAGGAGACTTTCAGATC 3482
Db 3440 CTGTGTTTAAAAATTTCAATCTTTTGTAGCAGCTCCAAAGACTGGAGATATATCTGATA 3499
Qy 3483 TACAGTTCTATTCAGATGTATGCTCCTGTTGTAAGTACTATATCTTAACAGATGATG 3542
Db 3500 TGCAATTTTACTATGTAAGTGTCTTCTGGGAAACAGCAGCTGTTGTGAACAACACTACG 3559
Qy 3543 CTGTTTACCATGAGGTTTACGTGATAATAGTCTTAATGTGAAGGATTTGTTCTTGAATTTT 3602
Db 3560 CTGTTTACCATGAATTTGACTGACATTTCTCTGAATGTCAAGATTTGCATATTAGATATG 3619
Qy 3603 CCAAAAGTATTCGATGCAAGGAGGTGAACCAATGTCTAGAGCCAGTTTTCGTTACCG 3662
Db 3620 CTAAGTCTGTAGTCTCTCGAAAGATGTCAAACCAACTTTAATACCGATGGTACGAACG 3679
Qy 3663 CGGCGGAACCGCAAGGCTGCGAGCTACTCGAAAATCTGCTCAATGATTTAAAGAA 3722
Db 3680 CGGCAAGAAATGCTCCGACAGCTGAGCTGTGTGAAAATCTAGTTGCGATGATTAAGAA 3739
Qy 3723 ATTTCAAACGACACAGACCTGACGGGAGGATTTGACATTTAGAGACCGCATCTGTTGTAG 3782
Db 3740 ATTTTAAATTCACAGAGTTGTCCGGAGTAGTTGATATTGAAAATACTGTCATCTTTAGTGG 3799

QY 3783 TAGATAAGTTTTTGTAGAGCTATTTTATTAAGAAAGAAATAATACACAAAAAATATTGCTG 3842
Db TAGATAAGTTTTTGTAGAGTTTTTACTTTAGGAAAAAGAAACCAACAAAAATTTTT 3859
QY 3843 GAGTGATGACGAAGGATTCAATGATGAGATGGTTGGAAAAACAGGAAAGAGTACTATTGG 3902
Db CACTGTTTAGTAGAGAGTCTCTCAATAGGTGATAGCAAGCAAGCAAGTCACAATTG 3919
QY 3903 ACGACTTGGCTAACTACAAATTTTACAGATCTCGCGGCATCGATCAGTACAGACACATGA 3962
Db GTCAAGTTGGCCGATTTTGAATTTTGTGATCTTCAGCCGTTGATCAGTACAGGCATATGA 3979
QY 3963 TCAAGGCTCAACCAAAAAACAGAAATTTGACCTTTTCAATTCAGAATGAATACCTCTCTGC 4022
Db TTAAGGCGAACCGAGCAGAAATCTGGATCTGTCAATTCAGACAGAAATATCCAGCGTTGC 4039
QY 4023 AAACAATTTGTCTACCAATTCGAAGCAGATCAACGGTATTTTGG----CGGTTTTCTCAGAGC 4079
Db AAACGATTTGTATCAATTCAGAAATAATCAACGCAATATTGTTCTCTTTTCACTGAGC 4099
QY 4080 TTACAAGGTTGCTGCTCGAGGCATTTGATTTCTAAGAAAGTTTCTTTTCTTTACTAGGAAAA 4139
Db TTACAAGGCAATTAATTGACAGATTTGACTCAAGCAGATTTCTTGTCTTTACGAGAAAGA 4159
QY 4140 CTCCAGAACAGATTCAGAAATTTTCTCGATCTCGACTCGCAGCTTCCTATGATGCTGT 4199
Db CACCGGCTCAGATCGAAGATTTCTTCGGAGATCTAGACAGTCATGTCCTCAATGACATC 4219
QY 4200 TAGAACTGGATATTCTTAAGTATGATAAGTACAGAAACGAGTTTCATTGTGCTGTAGAT 4259
Db TTGAGTTGGATTTTCGAGTATGATTAAGTCTCAAAACGAGTTTCATTGTGCTGTGAT 4279
QY 4260 ATGAAATATGGAAGAAGTTGGTCTCAATGAGTTTGTGGCGAAGTGTGGAACAAGGCG 4319
Db ACGAAATCTGGAGAGACTGGGTCTCGAGGATTTCTTGGCAGAGTGTGGAACAAGGCG 4339
QY 4320 ACAGAAACAACCTTTGAGGATTTACATTTGCTGGAATCAAGACATGCTGTGGTATCAA 4379
Db ATAGGAAACAACCTCTCAAGAGATTTACATCTGCTGTATAAACCGTGTATTGTGTACAGA 4399
QY 4380 GGAAGCGGTGATGCTACTTTTCATCGCAATCTGTTATATAGCAGCTGTCTGG 4439
Db GAAAGATGTTGATTTTACAACTTTTATCGGTATATACCGTCATCTTTGCTGTGCTAG 4459
QY 4440 GTTCAATGTTTACCGATGGAAGAGTCTATAAGGTGCTTTTGTGGAGACGATTTCCGTTT 4499
Db CATCAATGCTCCCGATGGAAGAAATTTGATAAAGAGAGCTTCTCGCGAGATGACAGTTGT 4519
QY 4500 TGTATTTTCCAAAGGTTTGGATTTCCCTGACATTCAGTCAATGCTAATCTCATGTGA 4559
Db TGTACTTTTCCTAAGGGTTGTGAGTATCCGATATACAAACAGCTGCTAATCTAATGTGA 4579
QY 4560 ATTTTGAAGGCAAACTGTATAGAAGAGGTACGGTACTTTTGTGGTAGATACATCATAC 4619
Db ATTTTGAAGGCCAACTGTTCAGAGCAATATGGGTACTTCTCGCGAGGTACGTGATTC 4639
QY 4620 ACCATGATAGGGAGCAATAGTGTATTATGATCTCTTTGAAGTTGATCTCCAAACTTGGGG 4679
Db ATCAGATAGAGGTTGCATAGTATACTACGACCTCTTGAAGCTGATTTGGAACCTTGGTG 4699
QY 4680 CAAAACATATCAAGGATATGATCACTTAGAAGAGTTAAGGTTCTTTTGTGCGATGTTG 4739
Db CTAACACATCAAGGATTTGGATCAATTTGGAGGAGTTTCAAGAGATCCCTCTGTGATGTTG 4759
QY 4740 CTTGTTCCGCTCGGAACTGGTCTTAGGCTTTCCGAGCTGAACGCACTATCAAGAGG 4799
Db CTGAGTCGTT---GAAACAATTTGCGGTATTAACAACAATTTGACGACGCTGTTGGGAGG 4816
QY 4800 TTCAATAAAACCGGATTTGATGGTTGGTTGCTTTTAAATTTGTTTAAACAAATTTTGTGTG 4859
Db TTCAATAAAACCGCCACCTGGTTGGTTGTTTATAGAGTTTAAAGTTTAAAGTTTGTGAG 4876
QY 4860 ATAAATTTTATTTAGAACTTTGTTTAAATGCTGTAGTCTCAGAGATCTGTCAAA 4919

Db 4877 ATAAAGTTTTTGTAGAGTTTTTCTTGTATGCTCTAGTGTGTTAAAGGTAAGGTAAT 4936
QY 4920 ATTACGAGTTTCATTTGATCTTTTGGAAAAACAGATGAGTACTTCCGGCAATTCATGACTAAG 4979
Db 4937 ATTAATGAGTTTATCGATCTGTCAAGTCTGAGAAACTTCTCCCGTGCATGTTCAACGCT 4996
QY 4980 GTCAAGAGTCTTAGAATATCGACTGTGCGAAGAGTATGCTGTAGAAAGTATGATGCTT 5039
Db 4997 GTAAAGAGTGTATGTTTCAAGGTTTGAAGAGTATGTTGTTCAATGAAAAATGAAATCAT 5056
QY 5040 TCTGATCTAGATTTTACTTAAAGGTTTAAAGTTAGTTAGTAAAGAGGTTATGCTGTAGCT 5099
Db 5057 TCTGAGTAAATCTCTTTAAAGGTTGTAATAACTTATAGAGTGGGTATGTTTGTAGTCT 5116
QY 5100 GATTGTTAGTGTCTGGGAGTGGAACTCTCCGGATAACTGCCCTGTGGTGTGCTCAGTGT 5159
Db 5117 GGTCTTGTGTGCTCGGTGAGTGGAAATTTACCAGATAATTTGCCGTGTGGTGTGAGTGT 5176
QY 5160 TGTATTTAGATAGAGAAATGAAAAAGAGTAAAGAAACAGCTGGGTGCGTATCACGCC 5219
Db 5177 TGCATGTTGACAGAGAAATGGAAGAGCGGACGAAGCCACACTGGGGTCTATATTACACT 5236
QY 5220 CCTGCTTGCAAAAAGAAATTTTCTTTTAAAGCTAATCCTTAATTTCAATAACATCCGAG 5279
Db 5237 GCTGCTCTAAAAAGCGGTTTCAGTTTAAAGTGTGCCAAACTACGGTATTTACAACAAG 5296
QY 5280 GATGCTGAGAGCAACCGCTGGCAAGTGTTAGTGAATATCAAGAGGAGTGGCTATGGAAGAA 5339
Db 5297 GATGCAAGAAAGACATATGCGAGTCTTAGTAATAATTAATAATGTAATAATGAGTGG 5356
QY 5340 GGATATGCTCTTTATCTTTGGAGTTCGTTTCAATTTGTGTAGTACATATAAAAAATATGTA 5399
Db 5357 GGTACTGCTCTTTGTCTATTAGAAATTTGTCTGTGTATTTGTTTATAAAAAATATA 5416
QY 5400 AGAAAGGTTTGGGGAACGTTATTTTGAAGTGTGACAGCGCTCGCAATTTGAATCTACT 5459
Db 5417 AATTGGGTTTGGAGGAGAAAGTAAACGAGTGTGAACGATGGAGACCCTCGAACTTTCA 5476
QY 5460 GAAAAGGTTTGTGAGGAGTTTCGTGGATGAAAGTACCAATGGCTGTGAACTCGAAAAAGTT 5519
Db 5477 GAAGAGTCTTGTGATGAGTTTCATGAGAAATGTTCCAAATGCGTTAGACTCGCAAGTTT 5536
QY 5520 CCGGAAAAACAAAAAGAAATGGTAGGTAAATAATGTTTAAAT----- 5561
Db 5537 CGAACCAAACTCTCAAAAAGAGGTTCCGAAAAATAATAATAATTTAGTGAAGGCGTTCA 5596
QY 5562 ---AAGAAATAATAACAGTGTGTAAGAGGTTTAAATAATTGAGGAAATTTGAGGATAAT 5618
Db 5597 GCGGGAAGGCTTAAACCAAAAAGTTTGTAGAAAGTTGAAAAGAGTTTGAATTTGAT 5656
QY 5619 GTAAAGTATGACGAGTCTATCGCGTCTATCGAGTACGTTTAAATCAATATGCTTTATACAA 5678
Db 5657 GAAGATGACGCGAGAGCTCGGTCCGGAATCTGATTCGATTAATAATATGTTCTTACTCAA 5716
QY 5679 TCAACTCTCGAGCCAAATTTGTTTACTTATCTTCCGTCTAGCAGATCTCTGTGAGCTGA 5738
Db 5717 TCACCTTCCCATCGCAATTTGTGTTTGTCTCATCTGTATGGGCTGACCCCTATAGAAATGT 5776
QY 5739 TCAATCTGTGTACAAATGCAATGGGTAAACGATTTCAACGCAACAGCTAGGACACAG 5798
Db 5777 TAAACGTTTGTACAAATTCGTTAGGTAAACGATTTCAACGCAACAGCAAGAACTACTG 5836
QY 5799 TCCAAACAGCAATTTGCGGATGCTCGGAAACCTGTGCTAGTATGACAGTGAATTTCTGT 5858
Db 5837 TTCAACAGAGATTTCAGCGAGGTGTGGAACCTTTCCCTCAGAGCACCGTCAATTTCTGT 5896
QY 5859 CATCGGATTTCTATGTGTATAGATATAATTCGACGCTTGATCCGTTGATCACGCGTTAT 5918
Db 5897 GCGATGTTTATAAGTGTACAGGTACAAATGAGTGTGAGTCTCTTAATTTACTGCGTTGC 5956
QY 5919 TAAATAGCTTTGATCTAGAAATAGAAATATAGAGTTTGAATTAATCAACCCCGACCGAATA 5978

Qy	2052	TG	CGAATGCTAACTTAGACACACAGAGATTGGAGTCCTCTCAACGATTTCCATAAAGGCTT	2111
Db	2059	CT	GAATCTTTATACACTAGGAGCGGAGAGATTGAGTCTCTCGACAGTTTTCATATGCGAA	2118
Qy	2112	CG	TGGATAGTGCATTACAAGCAAAATGGCATCGGTGTGTTCTACACATGGCTCACTCAAG	2171
Db	2119	CAG	CTAGTTCGTTAACTTAAAGCAGATGTGTTGATCGTGTAACGCGGCCCTCTTAAAG	2178
Qy	2172	TT	CAACAAATGAAGAACTATGTGCACAGTTTTGGCAGCTTCGTTCTCGCCACATGTATCAA	2231
Db	2179	TT	CAACAAATGAAGAACTTTATAGACAGCTGGTAGCTCGCTCTCTGCTCGGTGTCGA	2238
Qy	2232	AT	CTATGCAAGTCACTAAAGATGAAGTTCGGGTAATGATTCGTATTCAGGGAGAAAGTTG	2291
Db	2239	AT	CTAGTGAAGATCCTAAAGATACAGTCGCATCGACCTTGAAACTCGTCAAAAGTTCTG	2298
Qy	2292	GT	GTTCGGGATGTCACTTTGAAAGTGGCTCTCAAACTTCGCGCCAAAGGTCATTTTCAT	2351
Db	2299	GAG	TTCTGGATGTTCGTTTCGAAAGGTGGCTAGTTAAACCATTCGCGAAAGAACCATGCAT	2358
Qy	2352	GG	GAGTGTCTCGGATTAACAAGGGGAAAAATGTTTACTGCACTTCTATCTTATGAAGGAG	2411
Db	2359	GG	GGGGTGTGTGAGACTCATCGGAGGAATATACGTCGCACTTACTCGGACGATGAAT	2418
Qy	2412	AT	GAATGGTGACTGAGAGCGACTGAGGAGGGTGGCTGATATCATCTGATACAAATGGTAT	2471
Db	2419	TT	GCATTAATTACCTGCGATAACTGCGGACGCGGTGGCTGTGAGTCTGAGTCGGTAGTAT	2478
Qy	2472	ATT	CTGATATTGCAAAAGCTCCAAATCTGAGGAACAATGAGAGACGGTGAACCCACG	2531
Db	2479	ATT	CTGATATGGCTAAACTCAGGACTCTGAGAAGATTGCTCAAAAGATGSGAGAACCCACG	2538
Qy	2532	AAC	CTACTGCAAGATGGTACTTCTGTATGGGGTGCCTGGTTGTGGAAGAGTACAAAGGAG	2591
Db	2539	TT	AGTTCAGCAAGGTGGTTTTGTGTGATGGCGTTCCAGGGTCGGAAAAACGAGAGAA	2598
Qy	2592	AT	TTTGAAGATTTGATCTTGATGAGGATTTGATCTTTGGTTCCTCGGAAAAACAAGCTGCTG	2651
Db	2599	TT	CTTCGAGAGTTAAATTTGCAAGAAATCTAAATTTCTGTCCCTGGTCGTCAGCTGCCG	2658
Qy	2652	CT	ATGATCAGAAAGGGCTAAATTCATCTGGACTGTATAAGGCCACAAATCGACAAATGTGA	2711
Db	2659	AG	ATGATCAGAAAGAGACTAAATCGTCGGGCATTAATAGTGGCTACAAAGGATAAATGTGC	2718
Qy	2712	GA	CGGTAGATTCACTTCTAATGATCCAAAAACCGCATCAC-----ACAAAGAGGC	2762
Db	2719	GC	ACGTCGATCATTTTGTGATGAATTAACGGAAAGGGGACGCTGTGAGTTCAAAGAT	2778
Qy	2763	TTTT	TATGTAGAGGGTTGATGTCGACACCGGTTGTGTAACTTCCTCGTGCTTATCT	2822
Db	2779	TG	TTATAGACGAAAGTTGATGTCGATACTGTTGTGTGAAATTTCTTGGTTGAAATGT	2838
Qy	2823	CT	GGTTCGGACATCGCATACATTTACGGAGATACACAGCAGATTCCTTTCAATTAAACAGAG	2882
Db	2839	CT	TCAGCGGATTTGCGATATGTTTATGGAGACACCCCAACAAATTCGATCATCAACACAG	2898
Qy	2883	TT	CAGAAATTTCCCGTATCCCAACAATTTGAGAAAGCTGCAAGTGGATGGAAGTTGAGATGA	2942
Db	2899	TA	ACTGGTTTTCCCGTACCTGTGCACATTTGCAAAATTTGGAGGTGCGACGAAGTCGAAACAA	2958
Qy	2943	GG	AGGACCAACATGAGATGCCCGGTTGATGTGAATTTTTTTCCTACAAATCGAAGTACGAAG	3002
Db	2959	GA	AAACTACTCTTCGTTGTCCGCTGATGTGCACACATTTCCCTAAATCAAAAGGTATGAAG	3018
Qy	3003	GAG	CGGTGACAAACAATTCAACTGTACAAAGTCGGTCTCATCTGAGATGATAGGCGGTA	3062
Db	3019	GAC	ACGTAATGTGACGCTCTCTGAAAAGAAATCGGTTTCCCGAGAAATGGTTAGTGGGG	3078
Qy	3063	AG	GAGTACTAAACAGTGTTCCTCAACACACTTAAAGGGAAAAATTTGTAATCTTTCACCTCAGG	3122
Db	3079	CT	GGCTCTATCAATCTCTGTCTCAAGCGCTTAAAGGAAAAATTTTGACCTTTCACACAGT	3138
Qy	3123	CT	GATAAATTTGATTAGAGGAAAGGGCTATAAGAAATGTGAACACACCGTTTCATGAGATCC	3182

[illegible]

Db 4219 TTGAGTTGGATGTTTCGAAGTATGATAAGTCTCAAAACGAGTTTCATTGTGCTGTTGAGT 4278
Qy 4260 ATGAATAATGAAAGAGTTGGTCTCAATGAGTTTGGCCGAGTGTGGAACACAGGCG 4319
Db 4279 ACGAAATCTGGAGGAGACTGGGTCTGGAGGATTTCTTTGGCAGAGTGTGGAACACAGGCG 4338
Qy 4320 ACAGGAAACCACTTTGAAAGGATTTACATTCCTGGAATCAAGACATGCTGTGATCAAA 4379
Db 4339 ATAGGAAACCACTCTGAAAGATTAACATCTGCTGGTATAAACAGTGTATGATCCAGA 4398
Qy 4380 GGAAGAGCGGTGATGTAATCTTTTCATCGGCCAATCTGTTATAATAGCAGCTTGCTTGG 4439
Db 4399 GAAAGAGTGTGATGTTTACAACTTTTATCGGTAATACCGTCACTATGCTTCGTGCTTG 4458
Qy 4440 GTTCAATGTTACCGATGGAAGAGTCTATAAGGTGCTTTTGTGGAGACATTCGCTTT 4499
Db 4459 CATCAATGCTCCGATGGAAGAAATTTGATTAAGGAGCTTTCTCGGAGATGACAGTTTGT 4518
Qy 4500 TGTATTTTCCAAAGGTTTCGATTTCCCTGCACATTCAGTCAATGCTCAATCTCATGTGA 4559
Db 4519 TGTACTTTCTTAAGGTTGTGAGTATCCCGATATACAAAGCTGCTCAATCTAATGTGA 4578
Qy 4560 ATTTTGAGGCCAACTGTATAGAAAGAGGTACGCTTACTTTTGTGATAGATACATCATAC 4619
Db 4579 ATTTTGAGGCCAACTGTTTAAGAGCAATATGGTACTTCTCGGAGGTAAGTATTC 4638
Qy 4620 ACCATGATAGGAGCAATAGTGTATATGATCTTTGAAAGTGTATCTCAAACTTGGGG 4679
Db 4639 ATCAGATAGAGGTGTCATAGTATCTACGACCTTTTGAAGCTGATTTCCGAACTTGGTG 4698
Qy 4680 CAAAACATATCAAGGATTAATGATCATTAGAACGTTAAGGTGCTTTTGTGCGATGTTG 4739
Db 4699 CTAAACACATCAAGGATTTGGATCAATTTGGAGGATTTCAAGATCCCTCTGTAATGTTG 4758
Qy 4740 CTTGTTGCTCGGAAACTGCTGCTTAGGCTTTTCGACGCTGAAACGAGCTATCAAGGAGG 4799
Db 4759 CTGAGTCGTI--GAACAAATGCGCGTATTACACAAATTTGGACGCTGTTGGGAGG 4815
Qy 4800 TCCATAAACCGGATGATGTTGTTGTTGTTTAAATGCTGTAGTCTCAGAGATCTGTCAAA 4859
Db 4816 TCCATAAACCGGCCACCTGTTGTTGTTGTTTAAAGATTTAGTATTAGTATTGTTGTCAG 4875
Qy 4860 ATAAATTTTATTAGAACTTTGTTTAAATGCTGTAGTCTCAGAGATCTGTCAAA 4919
Db 4876 ATAAAGTTTGTGTAGAAATTTATTTCTGATGCTCTAGTGTGTTAAAGGTAAGTAAAT 4935
Qy 4920 ATTAGCGAGTTCATTGATCTTTGAAACAGGATGAGATCTTCGGCAATTCATGACTAAG 4979
Db 4936 ATTAATGAGTTTATCGATCTGTCAGAGTCTGAGAACTTCTCCGTCGATGTTCAAGCCT 4995
Qy 4980 GTCAGAGTGTAGAAATPATCGACTGTGGAACAAGATTAGCTGTTAAGAAATGATGCTTT 5039
Db 4996 GTAAAGAGTGTATGTTTCAAGGTTGTAAGATTTATGCTCCATGAAATGAATCATTG 5055
Qy 5040 TCTGATGTAGATTTACTTAAGGTTTAAAGTGTAGTTAAGAAAGGATGTTGCTTAGCT 5099
Db 5056 TCTGAAGTAAATCTTTTAAAGGTTAAACTTATAGAAGGTGGGTATGTTTGGCTTAGTT 5115
Qy 5100 GATTTGTGATGCTGTCGGGAGTGAATCTCCCGATAATTCGCGTGGTGTGTCAGTGT 5159
Db 5116 GGTCTTGTGTCGGGTGAGTGGAAATTAACAGATTAATTTCCGCTGGTGTGATGATGTC 5175
Qy 5160 TGTATTTGTAGATAAGAGATGAAGAGGATGAAGAACACGCTGGGTGCTATCAAGCC 5219
Db 5176 TGCTTGGTTGACAAGAAATGGAAGAGCGGACGAGCCACACTCGGGGTCAATATCACT 5235
Qy 5220 CTTGCTGCAAAAGAAATTTTCTTTTAAGCTAATCCCTAATTTATCAATAACATCCGAG 5279
Db 5236 GCTGCTGCTAAAAGCGGTTTCAGTTTAAAGTGTGCTCCCAATTAACGATATTACAACAAAG 5295
Qy 5280 GATGCTGAGAGACACCGTGGCAAGTGTAGTGAATATCAAGAGGATGGCTATGGAAGAA 5339
Db 5296 GATGCAAGAAAGAACATATGCGAGGCTTTAGTAAATATTAAATATGTAAGATGAGTGCG 5355

Qy 5340 GGATACGTGCTTTATCTTTGGAGTTCGTTTCAATTTGTGTAGTACATAAAATAATGTA 5399
Db 5356 GGCTACTGCGCTTTGTTCATTAGAAATTTGTGTCTGTGTATTTATATAAAATAATACA 5415
Qy 5400 AGAAAAAGTTTGAAGGAAACGTATTTTGAGTGTGACAGACGCTCGCCAAATTTGAACCTCACT 5459
Db 5416 AATTGGTTTGAAGGAGAAAGTAAACGAGTGTGAAACGATGAGGAGCCATCGAACTTCA 5475
Qy 5460 GAAAAAGTTTGAAGGAGTTCGTGGATGAAGTAAACCAATGGCTGTGAAACTCGAAAAAGTT 5519
Db 5476 GAAGAAAGTTTGTGATGAGTTTCATGAGTAATGTTCCAAATGTCGTTAGCTCGCAAGTGT 5535
Qy 5520 CCGGAAAAAAGAAATGGTAGTAAATGTTAAATATTTAGGTAAGGGGCGTTC 5561
Db 5536 CGAACCAATCTCAAAAAAGAGTCCGAAAAATAATATAATTTAGGTAAGGGGCGTTC 5595
Qy 5562 ---AAGAAAAATAATAACAGTGTGAAGAGGTTTTAAAAATTTGAGGAAATTTGAGGATAAT 5618
Db 5596 GCGGAAAGGCTAAACCAAAAGTTTTTGATGAAGTTGAAAAAGAGTTTGATATAATTTGATT 5655
Qy 5619 GTAAGTATGAACGAGTCTATCGCGTCAATCGAGTACGTTTTTAAATCAATATGCTTATCAA 5678
Db 5656 GAAGATGAAGCGAGAGCTCGGTGCGGATTTCTGATTCGTTATAATATGCTTACTCAA 5715
Qy 5679 TCAACTCTCGAGCCAAATTTGTTTACTTATCTTCGCTTACGAGATCTCTGTGCGAGCTGA 5738
Db 5716 TCACTTCTCCATCGCAATTTGTTGTTTGTCTGTATGGGCTGACCTTATAGAATGT 5775
Qy 5739 TCAATCTGTGTACAAATGCAATTCGGTAAACCAAGTTTCAACGCAACAGCTAGGACAAACAG 5798
Db 5776 TAAACGTTGTATCAAAATTCGTTAGGTAAACAGTTTCAACAGCAAGCAAGCACTACTG 5835
Qy 5799 TCAACAGCAATTTGCGGATGCTGAAACCTGCTGCTAGTATGACAGTGAAGTTTCTG 5858
Db 5836 TCAACAGCAGTTTCAGCGAGTGTGGAACCTTTCCCTCAGAGCAGCTCAGATTTCTG 5895
Qy 5859 CATCGAATTTCTATGTATAGATATAATTCGAGGCTTGATCCGTTGATCACGGCGTTAT 5918
Db 5896 GCGATGTTTATAAGGTGTACAGGTACAATGCAATTTAGATCTCTTAATTTACTGCGTTC 5955
Qy 5919 TAAATAGCTTTGATATCTAGAAATAGAAATATAGAGTTGATTAATCAACCCGCAACCAATA 5978
Db 5956 TGGGCTTTCTGATCTAGGATAGAAATATCGAGTGAAGAACACAGCAGAGTCCGACAA 6015
Qy 5979 CTACTGAAATCGTTAAACGAGCTCAGAGGTAGACGATGCTACTGTAGCTATAGGCTT 6038
Db 6016 CAGCTGAAACGTTAGATGCTACCGCAGGGTAGACGAGCTACGTTGCAATTCGGTCTG 6075
Qy 6039 CAATCAATTTTGGCTTAATGAATCGTTGCTGGAACCAACTCGGCTACTTACTTGTGTAG 6098
Db 6076 CTATAAATAATTTAGTTAATGAATAGTAAAGGTTACTGGAATGTTACAAATCAGAAATCTT 6135
Qy 6099 TTCAGAGCTGCTAGTGGACTTGTCTGGACCAAACTCGGCTACTTACTTGTGTAG 6158
Db 6136 TTGAAAGTATGCTGGGTTGCTGGACCTCTGCACTGCACTGATCTTAAATGATAGTGGC 6195
Qy 6159 ATTTCTTAAATAAAGTCTGCTGAAGACTTAAATAATTCAGGGTGGCTGATACCAAAATCAGC 6218
Db 6196 GAAATATAA-----GTTTGTGTTTCTTAAACACACGCTGCTACGATACCGTAC 6247
Qy 6219 AGTGGTTGTTGTCCTAATTAAATATTAACGATTTGTCATATCTGGATCCAAACAGTTAAACCA 6278
Db 6248 AGTGTGTTTCCCTCCACTTAAATCGAAGGGTGTGCTTGGAGCGCGGAGTAAACATA 6307
Qy 6279 TGTGATGTTGATATCTGTTGATGCGGTAAAAACATCGAGAGGTTTCAATCTCCCTAA 6338
Db 6308 TATGGTTTCAATATGTCGTTAGGACATGTAAGGAGGATTCGATTTCCCTCCCGAA 6367
Qy 6339 CCGCCGTTAGCGGCCCA 6355
Db 6368 CCGCCGTTGGGGCCCA 6384

RESULT 9
TMO243571
LOCUS
DEFINITION
Tobacco mosaic virus (K1 strain) complete RNA genome.
ACCESSION
AJ243571.1
VERSION
GI:6572455
KEYWORDS
126K protein; 17.7K protein; 183K protein; 29K protein; coat protein gene; movement protein; ORF1; ORF2; ORF3; RNA-dependent RNA polymerase.
SOURCE
Tobacco mosaic virus
ORGANISM
Tobacco mosaic virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE
1
Belenovich, E.V., Novikov, V.K. and Zavriev, S.K.
Biological properties and genome structure of the Kazakh strain K1 of tobacco mosaic virus
Unpublished
2 (bases 1 to 6383)
Zavriev, S.K.
Direct Submission
Submitted (19-JUL-1999) Zavriev S.K., Laboratory of Molecular Virology, Institute of Agricultural Biotechnology, Timiryazevskaya st. 42, Moscow 127550, RUSSIA
JOURNAL
Related Sequences: AF062519, Z92909.
FEATURES
Location/Qualifiers
1..6383
/organism="Tobacco mosaic virus"
/region
/variation
/mol_type="genomic RNA"
/strain="Kazakh strain"
/isolate="K1"
/db_xref="taxon:12242"
/country="kazakhstan"
join(72..3418,3422..4922)
/function="RNA replication"
/note="ribosomal slippage ORF1"
/codon_start=1
/product="rna-dependent RNA polymerase, 183K protein"
/protein_id="CAB62911.1"
/db_xref="GI:6572456"
/db_xref="GOA:Q9Q1T8"
/db_xref="Swiss-Prot:Q9Q1T8"
/translation="MAYQTATSSALLETVRGNTLVNDLAKRRLYDTAVDEFNARD RPKNVSKVSEETLIATKAYPEFQITFYNTQNAVHSLAGLSLELYLMQIPIYG SLTYDIGNFASHLPFKGRVYVHCMPNLDVRDIRHGGOKDSIELYLSRLERGNKVP NFOKEADRYAEMNEVYCHDTFOTCHRSQECYTRGVYVIAIALHSYDIPADEFGAALL RKNVHCYAAAFHSENLLEDSHVNLDEINACFORDDRLTFSPASESTLVNHSYSN ILKTYCTYFPASNREYVMEFLVTRVNTWFEKSRIDTFLLYKGVAAHGVNDVEQFYK AMEDAMHYKTLLAMCNSRIILLEDSSSVYFPMKMDVIPLFDSISLETSKRTRKEV LVSKDFVTVLNIHRTYQAKALTVSNLVSFVRSRVIINGVTARSEWDVKSLLQS LSWTFFLHTKLAVLDKDLILSKFALGPKTVSQHWDEIRISLAFGNAPPSIKERLINRKL IKITENALEIRVPDLYVTHDRLVSEYKMSDMPVLDIRKMEETEEMYNALSLSVL EKASDGLAVTSRDEEPIRSQMARGELQAGLSDVPESSTYRSEIESLEQFHW TASSLIHQMCSIYVTGPKVQOMKVFIDSLVASLSAAVSNLVKILKDTAIDLETRQ FGVLVDVASKRWLVKPSAKNHAWGVETHARKHIVALLEHDEFGLITCDNRRVAVSS ESUVYSDMAKLTRLRLKDGEPHVSAAKVLVDVPGCGKTKTEILSRVNFEDLILV PGROAEMIRRRANASGLIIVATKDNVTVDSFLMNYGKGARCOFKRLFIDEGMLHTG CVNFLVEMSLCDIIVYGDYQOIQIPYINRVTFPPYPAHFAKLEVDEVEFTRITLRCPAD VTHFLNORYEGHVCMCTSEKKSQEMVSGAASINPVSKPLGKILITFTQSDKEALLS RGYADVTHVEHQGTAYADVSLVRLTPTFVSIIRARDSPHVLVSLSRHTKSLKYITVM DPLVSIIRDLERVSYLLDMYKVDAGTQ"
4906..5700
/gene="ORF3"
4906..5700
/gene="ORF3"
/gene="ORF3"
/function="TMV movement in plants"
/codon_start=1
/product="29K movement protein"
/protein_id="CAB62913.1"
/db_xref="GI:6572458"
/db_xref="GOA:P03584"
/db_xref="Swiss-Prot:P03584"
/translation="NALVVKGVNINEFIDLSKSEKLLPSPMTPVKVSMVSKVDKIMV HENESISEVNLKGVKLEGGVYCLGVVSGEWNLPDNCRGVSVCMVDKMERADE ATLGSYTTAAAKKRFQFKVVPYNGITTDAAEKNIQWLVNIKNVMSAGYCPLSLEFV SVCSYITAAKIKLGRKTVSNVGGPMELSEKSVNDEFMENVPMSVRLAKFRYKSKRG PRKNNNLKGKSGRGPKPSPDFEVEKEDNLEDSEATSVADSDSY"
5703..6182
/gene="coat protein"
5703..6182
/gene="coat protein"
/function="viral RNA coating"
/codon_start=1
/product="17.7K protein"
/protein_id="CAB62914.1"
/db_xref="GI:6572459"
/db_xref="GOA:Q9Q1T6"
/db_xref="Swiss-Prot:Q9Q1T6"
/translation="MSYPIPSQFVFLSSVWADPIELNVCTNSLGNQFOTQOARTI VQOQSEVMEFPQSTVRFPDGVYKYRYNAVLDPLITALLGFTFTRNRIIEVENQRQS PTTAETLIDATRRVDDATVAIRSAIINNLVNLVRLVGTGLYNQNTFESMSGLVWTSAPAS"
ORIGIN
Query Match 41.3%; Score 2623.8; DB 14; Length 6383;
Best Local Similarity 64.4%; Pred. No. 0;
Matches 4095; Conservative 0; Mismatches 2227; Indels 55; Gaps 8;
QY 18 CGACAAACAAATTTAAAAACAAAAACAACTATTATCAAAACAAAAACAAACAAATGGCAC 77
DB 23 CAACACACACACACACACACATTTACATTTTACATTC----TACAACTACAAATGGCAT 78
QY 78 ACATACAACTTATAATAGCAACGCCCTTCTTGAAGCGTGAGTGTGTAAACAACTCTCG 137
DB 79 ACACACAAACAGACACATCGTCGCTTTTGTGACACCGCCGAGGTAAACAACTCTTGG 138
DGSSC"

QY 138 TTAATGACCTTGCAGNAGCGGCATGTACGATACGCGCGTGGAGAAATTTAAACGCGCGG 197
Db 139 TCAACGACTTGTCAAAAGCGCGCTATATGACACAGCGGTCGATGAATTTAATGCTAGGG 198
QY 198 ACCGTAGAACCAAGGTCAACTTTTCCAAAATATTAGCGAAGAGCAACGCTTCTAGTCT 257
Db 199 ACCGAGGCTTAAGTCNAATTTTCCAAAGTAGTAGGAGAGAACAGACGCTTATTGCAA 258
QY 258 CCAACGCGTACCGGAGTCCGATTAACCTTTTATTAATCTCAAAATGCCGATACACAGTT 317
Db 259 CCAAGCCTACCCAGAAATCCAAATTTACATTTCTACAAACGCGAGAACGCTGTGCAATCCC 318
QY 318 TGCGTGGAGTTGAGAGCATTAAGATTTGAATATCTGATGCTACAGTTCCCTATGGAT 377
Db 319 TTGAGCGCGTCTCCGATCATTAAGAAATTTGAATATCTGATGATGCAAAATTTCCCTACGGAT 378
QY 378 CGCGGACATATGATATAGTGGGAACTTTGACGACATTTGTTTCAAAGGCGAGGATTAACG 437
Db 379 CATGACATATGATATCGAGGTAAATTTGCACTCTCATCTGTTCAAAGGCGAGCATACG 438
QY 438 TGCAATGCTGATGCCCAATCTGGACATACGAGATATAATGAGGCAAGAGCAAAAGG 497
Db 439 TTCACTGCTGATGCCGAATCTGGATGTCGCGACATATAATGCGGCACGAGGCGCAAAAGG 498
QY 498 ACTCAATGAGATGATTTGTCGAGATTTGCTGTTCTAAAGGTAATTTCTGAGTTTC 557
Db 499 ACAGTATTGAACATATACCTTTTAGGCTCGAGAGGGGCAACAAACATGTCCCAAACTTC 558
QY 558 AAAGGAGGCTTTTAAACAGGTATGCGAAGCTCCCAACGAAAGTCTGCTGCTCTAAAACAT 617
Db 559 AAAGGAGGCTTTGACAGATACGCTGAAATGCGAAGAGTAGTCTGTACAGTACTT 618
QY 618 TTCAGGATGTCGAATACATCCGCGAGAGATAGTGTAGAGATAGCTGTTGCTCTGC 677
Db 619 TCCAAACGCTAGGCATTTCTCAAGATTTTACACGGGAAGGTATGCTATTGCTTTGC 678
QY 678 ACAGTTGTATGATATTCCTGTGATGAGTTTGGAGCTGGTTAATATCTAAGAAATATAC 737
Db 679 ATAGTATATACGATATACCTCGCGAGAGTTGCGCGCGGCACTGCTGAGAAAGAAATGTAC 738
QY 738 ATGTATGTATGACGCTTCCATTTTGGCAGAAGCATTTACTTAGACCAAGCGAGGTTA 797
Db 739 ATGTATGTATGCGCTTTCCACTTTTCCGAGAAATTTACTTCTCGAAGATTCACAGTCA 798
QY 798 CGCTTAATGAATAGGCGCACTTTTCAAAGAGAGGTGATGATGTTTCTTTTCTTTG 857
Db 799 ACCTCGATGAGATCAATGCAATGTTTCCAAAGAGATGGAGACAGGTTGACTTTTTCTTTG 858
QY 858 CTGATGAAGTACTTTAAATATATAGTCAATAAATACAAAATATCTTGCAATATGTAGTTA 917
Db 859 CATCTGAGAGTACTTTAAATATATAGTCATAGTTAATCTTAATTTCTAAGTATGTTGCA 918
QY 918 AATCTTACTTTCTGCTCTAGTAGAATAGTTTACTTTTAAGGAATTTTATGCTCACTAGGG 977
Db 919 AAATCTTCTCCAGCCTCTAAATAGAGAGGTTTACATGAAGGAGTTTATGTAACATAGAG 978
QY 978 TTAATACTTGGTTGTGAAATTTTACCAAGTAGATACCTATATTTCTGTAACAAGAGTTA 1037
Db 979 TAAATACCTGTTTGTAAATTTTCTAGATAGATACCTTTCTTATTGTACAAGGTGTAG 1038
QY 1038 GACAAGTAGGTTGATGATGATGATCTATGAGGCGATGGAAGAGCGCTTTGCTTTACA 1097
Db 1039 CGCATAGGGGTGAGACAATGAGCAGTTTTCAAAGGCTATGGAAGAGCGCATGGCACTACA 1098
QY 1098 AGAAACCTTGGCCATGTTCAACACTGAAGAGCAATCTTTTAGAGACAGCGCTTCGTTA 1157
Db 1099 AAAGACTCTTGGCATGTGCAACAGTGAAGAAATCTTTGTTAGAGGATTTCTCATCAGTTA 1158
QY 1158 ACTTTTGGTTCCCTAAGATGAAGGACATGGTGTATGATACCGCTGTTTGGGGTTCTATTA 1217
Db 1159 ATTACTGTTTCCAAAATAGGGATATGTTGATAGTTCCACTATTTGACATATCTCTCG 1218

QY 1218 CCAGCAAAAGAGATGCAAGGAGTGAGGTCAATCTTAAATCGTGACTTCGTTTACACAGTGC 1277
Db 1219 AGACTAGTAAAGAACACGAAAGAGGTCTTAGTTTCAAAGGACTTTGTTTATACAGTGT 1278
QY 1278 TTAATCATATCAGAAATATCAAGCCAAAGCGTTAACTTACCAAGAACGTATATCTTTG 1337
Db 1279 TAAATCACATTCGTACGTACCGGCCAAAGCGCTTACTTACTCCAAGCTGTATCTTTG 1338
QY 1338 TGGAGTCTATAGATCCCGGTGATTAATCAATGGTGTACTGTAGTCTGATGGGATG 1397
Db 1339 TCGAATCAATTCGTTCCGAGAGTGATCAITTAACCGGGGTACTGTAGTCTGAGTGGGATG 1398
QY 1398 TAGATAAGCAATTTCTTCAACCTCTCAATGACTTTCTTCTTGGACACTTAAGCTGCTG 1457
Db 1399 TCGATAAATCATTAATTAACAGTCTCTTGTGATGAGCTTTCTTCTACATACCAAGCTTCCG 1458
QY 1458 CGCTTCAAGACGATATAGTAATGGGAAAGTTTCGGTCTTTGGATAAGACCACTTCTGAAC 1517
Db 1459 TTCTGAAGACGATCTTTTGAATTAGCAAGTTTGCACCTTGGTCCAAAACCTGTCTCAAC 1518
QY 1518 TTAATTTGGGATGAGGTGGGCAAAATTTTGGAAAACGTTTTCGCCACTATCAAGAGAGAT 1577
Db 1519 ATGTGTGGGATGAGATTTCCCTAGCTTTCGGCAATGCTTTCCCATCGATCAAGGAAAGAT 1578
QY 1578 TGGTGAGCAGGAAAATTTCTGGATGTAAGTGAGAAATGCTCTGAAGATCAAGATCCAGATC 1637
Db 1579 TGATAAACCGAATCTGATCAAAATTAACGAGAAATGCGTTAGAGATCAGGGTCCCGATC 1638
QY 1638 TGTATGTCATGTAAGAGACAGGTTCTGTAGCTGAATACACAAAGTCTGAGGAGTTACCGC 1697
Db 1639 TTTATGTCACCTTCCATGATAGGTTAGTTTCTGAGTACAAAATGTCAAGTGCATGCGG 1698
QY 1698 ATCTAGATATCAAGAAAGACTTAGAAGAGCTGAGCAAAATGTACGAGCGGTTATCAGAA 1757
Db 1699 TGCTAGACATATTAGAAAAGATGGAAGAACTGAGGAAATGTACAAATGTCACTGTCCGAA 1758
QY 1758 TATCTATCTTAAAGGTGCTGAATAATTCGATATCCGAAATTCAGAGATGTCGAAGG 1817
Db 1759 TGTCTGTACTTAAACCTCAGACAGTTCGATGTTGATGTTTTCCTCCAGATGTCGAAT 1818
QY 1818 CTTTATAGTGTAGTCTGTGCGCAGCAGAGTAATCGTTGCGAGTGGCGCGAGAAATAGAA 1877
Db 1819 CTTTATAGTTCGATCCATGACTGCGACGAAAGTATAGTAGCAGTTATGAGCAACGAGA 1878
QY 1878 GCGTTTAACTCTTACTTTTATGATAAGCAACCGAGGAGAAATGTGCGTAAAGGCTTTAAA 1936
Db 1879 GCGGTCTTACTCTCAGCTTGAACAGCCACCAAGCTAATGTTGCGCTAGCATTTGCAAG 1938
QY 1937 -----AGCAGCGGCTGAGCGCGTGGTATGTTCTTGAACCGACATCCGAGAGGTGAACG 1991
Db 1939 ATTCTGAAGGGCTTCTGACGGGCGGTTGGTAGTTTACCTCAAGAGATGTTGAGGAACCGT 1998
QY 1992 TAAATAAATTTTCTATTGCTGAGAAAGGAGATTTGCTGTGTGTCAGAAAGTCAATGGTT 2051
Db 1999 CCATAGGGGTTGATGCGCGGCTGGTGAGTTACAAATGGCGCGGATTTCTGGCGACGTT 2058
QY 2052 TGACGAATGCTAATCTTATAGACACAGAGTTGAGTCCCTCAACGATTTCCATAAGGCTT 2111
Db 2059 CTGAATCTTCAACACTAGGAGCGAGAGATTGAGTCTCTCGAGCAGTTTTCATATGCA 2118
QY 2112 GCGTGTAGTGTGATTAACAAAGCAATGCGATCGGTTGCTACACTGGCTCACTCAAG 2171
Db 2119 CAGCTAGTTGTTTAAATTAAGCAGATGTTTTCGATCGGTGATACCGGGCCCTCTTAAG 2178
QY 2172 TTCAACAAATTAAGAACTATGTGACAGTTTGGCAGCTTCTGTTGTCCGCACTGTATCAA 2231
Db 2179 TTCAACAAATTAAGAACTTTATAGACAGCTGCTGCTAGCTCTCTGCTGCGGTGCGA 2238
QY 2232 ATCTATGCAAGTCACTAAAGGATGAAGTGGGTATGATTTCTGATTTCCAGGAGGAAAGTTG 2291
Db 2239 ATCTAGTGAAGATCTTAAAGATACAGCTGCGATGACCTTGAACCTCGTCAAAAGTTG 2298
QY 2292 GTGTTTGGGATGTCACTTTTGAAAAAGTGGGCTCTCTCAACCTGCGGCCAAAGGTCATTCAT 2351

Db 2299 GAGTCTCGATGTTCTTCGAAAAGGTGGCTAGTTAAACCAATCCGCAAGAACCAATGCAT 2358
Qy 2352 GGGAGTTGCTCTGATTAACAAGGGGAAAATGTTTACTGCACTTCTATCTATTAAGAGAG 2411
Db 2359 GGGGGTGTGTGAGACTCATCGAGAGAAATATACGTCGCATTAATCGGACACGATGAAT 2418
Qy 2412 ATAGAATCGTGACTGAGAGCACTGGAGAGGGTGGCTGTATCATCTGTATACAATGGTAT 2471
Db 2419 TTGGCAATTAACGTGGCAACACTGGCGACGGGTGGCTGTGAGTCTGAGTCCGTAGTAT 2478
Qy 2472 ATTCTGATTAATGCAAAAGCTCCAAATCTGAGGAAAACAATGAGAGAGCGGTGAACCCACG 2531
Db 2479 ATTCTGATATGCTAAACTCAGGACTCTGAGAGAAATGCTCAAAAGATGGAGAACCAACG 2538
Qy 2532 AACCTACTGCAAAAGATGTTCTTGGATGGGTGGCTGGTGTGTTGGAAAGTACAAAGGAG 2591
Db 2539 TTAGTTACGAAAGGTGGTGTGTGGATGGCTTCAGAGGTGGCGGAAAACAAGGAAA 2598
Qy 2592 ATTTTGAAGATTTGATCTTGATGAGGATTTGATCTTGGTTCTGGAAAACAAGCTGCTG 2651
Db 2599 TTCTTTGAGAGTTAAATTTGAGAGAGATCTAATTTCTGTCCCGGTGGTCAAGCTGCG 2658
Qy 2652 CTATGATCAGAAAGGGCTAATTCATCTGCACTGTGATGATAAGAGCCACAATGACAATGTGA 2711
Db 2659 AGATGATCAGAAAGAGCTAATGCTCGGCATAATAGTGGCTTACAAAGGATAATGTGC 2718
Qy 2712 GAACGGTAGATTCACCTTCTAATGATCCAAACCGCGATCAC-----ACAAAGGC 2762
Db 2719 GCACGTCGATTCATTTTGTATGAATTAACGGAAAGGGGCAAGCTGTGAGTTCAAAAGAT 2778
Qy 2763 TTTTATTTGATGAAGGGTGTGCTGCACACCGGTGTGTAACTTCTCGTGGTCTTATCT 2822
Db 2779 TGTTTCATAGACGAAGTTTGATGCTGCATCTGTTGTGTGAATTTCTTGGTTGAAATGT 2838
Qy 2823 CTGTTGGACATCGCATACATTTACGGAGATACACAGCAGATTCCTTTTCATTAACAGAG 2882
Db 2839 CTCTGTGCGATTTGTCATATGTTTATGGAGACACCCCAACAAATTCCTGTACATCAACAGAG 2898
Qy 2883 TTCAGAAATTTCCGTTATCCCAACATTTTGAGAGAGCTGCAAGTGATGAAGTTGAGATGA 2942
Db 2899 TAACTGGTTTCCCGTACCTGCGACACCTTTGCAAAATTTGGAGGTGCGCAAGTCCAAACAA 2958
Qy 2943 GGAGGACACACTGAGATGCCAGGTGATGTGAATTTTTTCTCAATCGAAGTACGAAG 3002
Db 2959 GAAGAACTACTCTTCGTTGTCCGCTGATGTCAACACACTTCTTAAATCAAAGGTATGAAG 3018
Qy 3003 GAGCGGTGACAAACCACTTCACTGTACAAAGATCGGTCTCATCTGAGATGATAGGCGGTA 3062
Db 3019 GACACGTAATGTGCACGTCTTCTGAAAGAAATCAGTTTCCAGGAAATGGTTAGTGGGG 3078
Qy 3063 AGGAGTACTAAACAGTGTTCCTCAACCACTTAAAGGGAATTTGTAACCTTCACTCAGG 3122
Db 3079 CTGGCTCTATCAATCTGTCTGCAAGCGCTTAAAGGGAATTTTGAACCTTTCACAGT 3138
Qy 3123 CTGATAAATTTGAGTTAGAGAGAGGGCTATAGAATGTGAACACCGTTTCATGAGATCC 3182
Db 3139 CTGAACAGGAGGCCCTTCTCTCAAGGGGCTATGAGATGTCCATCTGTACATGAGGTAC 3198
Qy 3183 AAGGAGAAACCTTTGAAGATGTGCTGCTGATGAGTTGACGGGCACTCCACTGACTCTGA 3242
Db 3199 AAGGTGAGACGTATGCGAGCGTATCGTTAGTTGAGCTTAAACCTTCAAGCCCTGTATCTATCA 3258
Qy 3243 TTTTCAAGTCTTCCCGCATGTTCTAGTCCGTCTGACTAGACACACAAGAGCTTCAAAAT 3302
Db 3259 TCGCAAGAGACAGTCCCGCATGTTCTGTCTCGTTGTCAAGACACACAATAATCCCTAAGT 3318
Qy 3303 ATTACCCGTAGTTAGATCCCTTTAGTACAGATAATGATGTTGCTCTTTTAAGCT 3362
Db 3319 ACTACCCGTGTGATGGATCCCTTTAGTTAGTATCAATTAGAGATTTAGAACGGGTAGTA 3378
Qy 3363 CCTTCCTTTTAGAAATCTATATGTTAGAACGAGGTAGTATAGCAATTAACAGATGATG 3422

Db 3379 GTTACTTTATTAGACATGTCACAAAGTAGATGCAGGTACTCAATAGCAATTAACAGTTCGACT 3438
Qy 3423 CAGTGTTCAAAGTCAATATCTCTTTGTGGCAACACTTAATCAGGAGACTTTTCCAGATC 3482
Db 3439 CTGTGTTTAAAAATTTCAATCTTTTGTAGCAGCTCCAAAGACTCGAGATATATCTGATA 3498
Qy 3483 TACAGTTCATTAACGATGATGCTCCCTGGTAATAGTACTATACATTTAAACAATGATG 3542
Db 3499 TGCATTTTACTATGATAGTGTCTTCTGGGAAACAGACGTTGTTGAACACTACGACG 3558
Qy 3543 CTGTTACCATGAGTTTACGTGATTAATGTGAAGGATTTGTTCTTTGATTTTT 3602
Db 3559 CTGTTACCATGAAATTTGACTGACATTTCTTGAATGTCAAAGATTCATATTAGATATGT 3618
Qy 3603 CCAAAAGTATTCGATGCCAAAGAGGTGAAACCATGTCTAGAGCCAGTTTTCGCTACCG 3662
Db 3619 CTAAGTCTGTAGTGTCTCCGAAAGATGTCAAACCAACTTTTAATACCAGTGTGTAACGCG 3678
Qy 3663 CGCGGAAACCGCCCAAGGGCTGCAGGACTACTCCGAAATCTGGTTGCAATGATTAAAAAGAA 3722
Db 3679 CGCAGAAATGCTTCGCCAGACTGGACTGTTGGAAATCTAGTTGCAATGATTAAAAAGAA 3738
Qy 3723 ATTTCAAACGACACAGACCTGACCGGACGATTCGATTCAGAGCACCGCATCTGTTGTAG 3782
Db 3739 ATTTTAATTCACCAGAGTTGTCCGAGTAGTTGATATTGAAAATACTGCATCTTTAGTGG 3798
Qy 3783 TAGATAAGTTTTTGTAGTATTTTAAABAAAGAAAATACACAAAATAATTGCTG 3842
Db 3799 TAGATAAGTTTTTGTAGTATTTTAAAGGAAAAGAAAACCAAAACAAAAATTTTT 3858
Qy 3843 GAGTGATGACGAAGATTCAAATGATGAGATGGTTGAAAACAGAAAAGAAAGTACTATTGG 3902
Db 3859 CACTGTTTGTAGAGAGTCTCTCAATAGTGGATAGCAAGCAAGCAAGTCAACATTTG 3918
Qy 3903 ACAGCTTGGCTAACTACAAATTTTACAGATCTGCGGCGCATTCGATCAGTACAAGCACATGA 3962
Db 3919 GTCAGTTGGCCGATTTTGTGATCTTCCAGCGTTGATCATCAGTACAGGCATATGA 3978
Qy 3963 TCAAGGCTCAACCAAAACAGAAATTCGACCTTCAATTCAGAAATGAATACCTGCTCTGC 4022
Db 3979 TTAAGCGCAACCGAAGCAGAAACTGGATCTGTCAATTCAGACAGAAATTCAGACGTTGC 4038
Qy 4023 AAACAAATTTGTACCAATTCGAAGCAGATCAACGGTATTTTGG---CCGGTTTCTTCAGAGC 4079
Db 4039 AAACGATTTGTATCATTTCAAAGAAATCAACGCAATATTGCTCTCTTTTCAGTGAGC 4098
Qy 4080 TTACAGGTTGCTGCTCGAGGCATTTGATTTCAAGAAAGTTTCTTTTCTTTTACTAGAGAAA 4139
Db 4099 TTACAAGGCAATTTACTTGAAGTATTTGACTCAAGCAGATTTCTTTTCTTTTACGAGAAAGA 4158
Qy 4140 CTCAGAACAGATTCAGAAATTTTCTCGGATCTCGACTCGCAGCTTCCTATGGATGTGT 4199
Db 4159 CACCGCTCAGATCGAAGATTTCTTCGGAGATCTAGACAGTCAATGTCCTCCANTGGAGTAC 4218
Qy 4200 TAGAACTGGAATTTCTAAGTATGATAGTCAAGAACGAGTTTCATTTGCTGTAGAGT 4259
Db 4219 TTGAGTTGAGTGTTCGAAGTATGACAGTCCCAAAACGAGTTTCATTTGCTGTGTAGT 4278
Qy 4260 ATGAATATGGAAGATTTGGGTCTCAATGAGTTTTCGCCGAGGTGTCGAAACAGAGGC 4319
Db 4279 AGAAATTCGAGGAGACTGGGTCTGGAGATTTCTTGGCAGAGGTGTCGAAACAGAGGC 4338
Qy 4320 ACAGAAAAACAATTTGAAGGATTAATTTGCTGGATCAAGACATGCTCTGCTGGTATCAAA 4379
Db 4339 ATAGAAAAACCACTCTGAAGATTAACATGCTGGTATAAAAAACGTTTATGGTACCAGA 4398
Qy 4380 GGAAGACGGTGATGAGTACTTTTCATCGGCAATCTGTTATAATPAGCAGCTTGTCTGG 4439
Db 4399 GAAAGAGTGGTATGTTTACAACTTTTATCGGTAATACCGTCACTGCTGCTGTCTG 4458
Qy 4440 GTTTCATGTTACCGATGGAAGAGGTATGAAAGGTGCTTTTGTGGAGAGGATTCGGTTT 4499
Db 4459 CATCAATGCTCCGATGGAAGAAATTTGATAAAGGAGCCTTCTCGGAGATGATAGTTGT 4518

Qy	4500	TGTAATTTTCCAAAGGGTTTGGAATTTCCCTGACATATCTAGTCATGTGCTAAATCTCATGTGGA	4559
Db	4519	TGTACTTTCTTAAGGGTTGTGAGTATCCCGATATACAACAAGCTGCTAAATCTAATGTGGA	4578
Qy	4560	ATTTTGAGGCCAAACTGCTATAGAAAGAGGTACGGTTACTTTTGTGTGTAGATACATCATAC	4619
Db	4579	ATTTTGAGGCCAAACTGTTCAAGAGCAATATGGGTACTTCTCGGGAGGTACGTGATTC	4638
Qy	4620	ACCATGATAAGGAGCAATAGTGTATATGATCCTTTTGAAGTTGATCTCCAAACTTTGGGG	4679
Db	4639	ATCACGATAGAGTTGCATAGTATACTACGACCCCTTTGAAGCTGATTTTCAAACTTTGGTG	4698
Qy	4680	CAAAAACATATCAAGGATATGATCACTTACGAAGTTAAGGGTGTCTTTTGTGCGATGTTG	4739
Db	4699	CTAAACACATCAAGGATTTGGGATCATTTGAGGAGTTTCAGAAGATCCCTCTGTGATGTTG	4758
Qy	4740	CTTTGTTCCGCTCGGAAACTGCTGCTTTAGGCTTTTCGGCAGCTTGAACGCAAGCTATCAAGGAGG	4799
Db	4759	CTGAGTCGTT---GAAACAATTCGCCGTATTACACACAATTTGGACGCGCTGTTCGGGAGG	4815
Qy	4800	TTCAATAAAACCGCGATTGATGTTGCTTTTAAATGTGTATAAATAATTTTGTGTG	4859
Db	4816	TTCAATAAAACCGCCACCTGGTTGCTTTGTTTATTAAGAGTTTAGTTAAAGTATTTGTCAG	4875
Qy	4860	ATAAATTTTTTATTAGAACCTTTGTTTTTAATATGGCTGTAGTCTCAGAGATACGTCAAA	4919
Db	4876	ATAAAGTTTTTGTATTAGAAAGTTAATTTCTGATGGCTCTAGTTGTTAAAGTAAGGTAAAT	4935
Qy	4920	ATTAGCGAGTTCATTGATCTTTTGGAAACAGATGAGATACTTCCGGCATTTCATGACTAAG	4979
Db	4936	ATTAAATGAGTTTATCGATCTGTCAAGTCTGAGAAACTTCTACCGTGCATGTTTCAAGCCT	4995
Qy	4980	GTCAGAGTGTTTGAATATACGACTGTGGAACAAGATTATAGCTGTTAAGAAATGATGCTTT	5039
Db	4996	GTAAGAGTGTTATGCTTTCAAAGGTTGATAAGATTATGCTCCATGAAATGAATCATTTG	5055
Qy	5040	TCTGATGTACATTTACTTTAAAGTGTTTAAAGTTAGTTAAGAAAGGGTATGTGCTTAGCT	5099
Db	5056	TCGGAAGTAAATCTCTTAAAGGTGTAAAACTTTATAGAAGTGGGTATGTTTCTGTAGTT	5115
Qy	5100	GATTTGGTAGTGTCTGGCGAGTGGAAATCTCCCGATAAATCTCCCGTGTGTGTGCTAGTGT	5159
Db	5116	GGCTTTGTTGTCGCGTGAGTGGAAATTTTACAGATAATTTGCCGTGCTGTGTGAGTGTCT	5175
Qy	5160	TGTATTGTGATTAAGAGAAATGAAAGAGGTAAAGGAAGCAACGCTGGGTGCGTATCACGCC	5219
Db	5176	TGCATGGTTGCACAAGAGATGGAAGACGCGAGCGAAGCCACACTGGGGTTCATATTACACT	5235
Qy	5220	CTTGCTTGCAAAAGAAATTTTCTTTTAAAGTAAATCCCTAAATTTATCAATAACATCCGAG	5279
Db	5236	GCTGCTCTTAAAGCCGGTTTCAGTTTAAAGTGGTCCCAATTCAGGTATTAACAACAAAG	5295
Qy	5280	GATGCTCAGAGAAGCACCGCTGGCAGTGTATTAGTGAATAATCAAGAGAGTGGCTATCGAAGAA	5339
Db	5296	GATGCAGAAAGAACAATATGCAGGCTCTTAGTAAATATTAATAATGTATAAAATGAGTGCG	5355
Qy	5340	GGATACGTGCTTTATCTTTTGGAGTTCGTTTTCAATTTGTGTAGTACATAAAAAATAATGTA	5399
Db	5356	GGCTACTGCCCCTTTGTCTATTAGAAATTTGTCTGTGTGTATGTTTTATAAAAAATAATATA	5415
Qy	5400	AGAAAAGGTTTGAGGGAACGTTATTTTCAGTGTGCACAGCGCTCGCCAAATTTGAACCTACCT	5459
Db	5416	AAATTTGGTTTGAGGGAGAAAGTAAACGAGTGTGAAACGATGAGGACCCATTTGAACTTTCA	5475
Qy	5460	GAAGAAGTTCTTGAGGAGTTCGTGGATGGAAGTACCAATGGCTGTGTAACCTCGAAAAGGTT	5519
Db	5476	GAAGAAAGTTGTGTAGTGTTCATGGAAGATGTTCCAAATGTCGGTTAGACTCGCAAGTTT	5535
Qy	5520	CCGAAAAACAAAAGAAATGTTAGTGAATTAATGTTTAATAT-----	5561
Db	5536	CGAACCAATTCCTCAAAAAGAGGTCCGAAAAATAATTAATTAATTTAGTGAAGGGCGGTTC	5595

QY	5562	---	AGAAATAAATAACAGTGGTGAACAGGGTTTAAATTTGAGGAAATTGAGGATAAT	5611
Db	5596	GGCGGAAGGCGCTAAACCAAAAGTTTGTATGAAGTTGANAAGAGTTTGTATAATTTTGATT	5655	
QY	5619	GTAAGTCATGACGAGTCTATCGCGTCACTGAGTACGTTTAAATCAATATGCCTTATACAA	5678	
Db	5656	GAAGATGAGCGCGAGAGCTCGGTTCGGGATTCGTATTCGATTAATAATATGTTTACCCAA	5715	
QY	5679	TCAACTCTCCGAGCCAAATTTGTTTACTTATCTTTCCGCTTACGCGAGATCCCTGTGCGAGCTGA	5738	
Db	5716	TCACTTCTCCATCGCAATTTGTGTTTGTCACTGTATGGGCTGACCCCTATAGAAATGT	5775	
QY	5739	TCAATCTGTGTACAAATTCGATTGGGTAACAGATTTCAAACGCCAACAGCTTAGGCAACAG	5798	
Db	5776	TAAACGTTTGTACAAATTCGTTAGGTAAACAGTTTCAAAACACAGCAAGCAAGCAACTACTG	5835	
QY	5799	TCCAAACAGCAATTTGCGGATCGCTTGGAAACCTGTGCCCTAGTATGACAGTGCAGATTTCCCTG	5858	
Db	5836	TTCAACAGCAGTTTACGCGAGGTGCGGAACCTTTCCCTCAGACACCGTCAGATTTCTCTG	5895	
QY	5859	CATCGGATTTCTATGTGTATAGATAATAATTCGACGCTTGATTCGGTTGATCACCGCGTTAT	5918	
Db	5896	GCGATGTTTATAAGTGCTATAGGTACAAATCGAGTTTGTAGATCCCTTTAAATTTACTGCGTTG	5955	
QY	5919	TAAATAGCTTTGATCTAGAAAAATAGAAATAATAGAGGTTTGATATCAACCCGCCACGAATA	5978	
Db	5956	TGGGCACTTTTCGATCTAGGAATAGAAATAATCGAAGTAGAAAAATCGGCAGAGTCCGACAA	6015	
QY	5979	CTACTGAAATCGTTTAAACGCACTCAGAGGCTAGACGATGCTACTGTAGCTATTAAGGCTT	6038	
Db	6016	CAGCTGAAACGTTAGATGCTACCCGCAAGGTAGACGACGCTACGGTTGCAATTCGGTCTG	6075	
QY	6039	CAATCAATAATTTGGCTTAATGAACTGGTTCGTTGGAATCGGCATGTTCAATCAAGCAGGCT	6098	
Db	6076	CTATAATAATTTAGTTTAATGAACTAGTAGAGGCTACTGGACTGTACAATCAGAATACTT	6135	
QY	6099	TTGAGACTGCTAGTGGACTTGTCTGGACCAACAATCCGGCTACTTACTGATTTGTTGTAG	6158	
Db	6136	TTGAAAGTATGCTCGGGTGGTCTGGACCTCTGCACCTGCATCTTAAATGCTAGATGCT	6195	
QY	6159	ATTTCCTAAATAAAGTCGCTGAAGACTTAAATTCAGGGTGGCTGATACCAAAATCAGC	6218	
Db	6196	GAATAATAA-----GTTTGTGTTTCAAACACACGCTGGTACGTACGATAACGTAC	6247	
QY	6219	AGTGGTTGTTCTGCCTCAATTAATAACGATTTGTCATATCTGATCCAAACAGTTAAACCA	6278	
Db	6248	AGTGTGTTTCCCTCCACTTAATTCGAAGGTAGTGTTCTGGAGCGCGGAGTAAACATA	6307	
QY	6279	TGTGATGGTGATATCTGTGTATGGCGTAAAAACATCGAGAGGTTGAAATCCCTCCCTAA	6338	
Db	6308	TATGGT-TCATATATGTCGTAGGACGTAAAAAAGGAGGGATTCGAATTCGCCCGGAA	6366	
QY	6339	CCGCCGTTAGGGGCCCA	6355	
Db	6367	CCGCCGTTAGGGGCCCA	6383	
RESULT 10				
AF155507				
LOCUS				
DEFINITION				
ACCESSION				
VERSION				
KEYWORDS				
SOURCE				
ORGANISM				
REFERENCE				
AUTHORS				
TITLE				
JOURNAL				
<p>1. (bases 1 to 6384)</p> <p>Bingsheng, Q., Gong, Y., Xiuhua, Z. and Po, T.</p> <p>Complete nucleotide sequence of attenuated tomato mosaic virus K genome</p> <p>Patent: China (00100211.2) 14-JAN-2000;</p> <p>CAS. Institute of Microbiology, Beijing, China</p>				
<p>AF155507</p> <p>Tobacco mosaic virus, complete genome.</p> <p>AF155507</p> <p>AF155507.2</p> <p>GI:5809676</p> <p>Tobacco mosaic virus</p> <p>Tobacco mosaic virus</p> <p>Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.</p>				
<p>6384 bp</p> <p>RNA</p> <p>linear</p> <p>VRL 02-MAR-2000</p>				

RESIST 10

AF155507
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
REFERENCE
AUTHORS
TITLE
JOURNAL

AF155507
Tobacco mosaic virus, complete genome.
AF155507
AF155507.2 GI:5809676
Tobacco mosaic virus
Tobacco mosaic virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
1 (bases 1 to 6384)
Bingsheng, Q., Gong, Y., Xiuhua, Z. and Po, T.
Complete nucleotide sequence of attenuated tomato mosaic virus K
genome
Patent: China (00100211.2) 14-JAN-2000;
CAS, Institute of Microbiology, Beijing, China

linear VRL 02-MAR-2000

REFERENCE 2 (bases 1 to 6384)
 TITLE Sheng, Q.B. and Gong, Y.
 JOURNAL Direct Submission
 Submitted (01-JUN-1999) Molecular Virology and Genetic Engineering,
 Institute of Microbiology, Chinese Academy of Sciences,
 Zhongguancun Beiyitiao No.13, Haidian District, Beijing 100080,
 People's Republic of China

REFERENCE 3 (bases 1 to 6384)
 TITLE Sheng, Q.B. and Gong, Y.
 JOURNAL Direct Submission
 Submitted (31-AUG-1999) Molecular Virology and Genetic Engineering,
 Institute of Microbiology, Chinese Academy of Sciences,
 Zhongguancun Beiyitiao No.13, Haidian District, Beijing 100080,
 People's Republic of China

REMARK
 COMMENT Sequence update by submitter
 FEATURES On Aug 31, 1999 this sequence version replaced gi:5524641.
 source
 1. .6384
 /organism="Tobacco mosaic virus"
 /mol_type="genomic RNA"
 /db_xref="taxon:12242"
 /note="attenuated mutant strain originally derived from
 nitrite-treated tomato strain"
 72. .4922
 /gene="replicase"
 72. .4922
 /gene="replicase"
 /function="genomic replication"
 /note="read-through protein"
 /codon_start=1
 /transl_except=(pos:2670..2672,aa:OTHER)
 /transl_except=(pos:3420..3422,aa:OTHER)
 /product="183 kDa protein"
 /protein_id="AAD44327.1"
 /db_xref="GI:5524644"

gene
 CDS
 /translation="MAYTQTATSSALLETVRGNTLVNDLAKRRLVDAADVFNARDR
 RPKVNFSEVSEQTLIATKAYPEFQITFYNTQNAVHSLAGLSLELYMMQIPYV
 SLTYDIGNFASHLFGKRAYVHCCMPNLDVDMRHEGQKDSIELYLSRERGNKHP
 NFOKEAFDAEMNEVWCHDTQTCTHSEOECTGRVYATALHSIYDIPADEFGAALL
 RKNVHCYAAHSENLLEDHVNDEINACFORQDRLTFSEFASSTLYNSHSYN
 ILKYCKTTPASNRVYKKEFLVTRVNTWFKESRIDTFLYKGVHAKGVDSQOFYK
 TMEDAHYKTKLWYNSERILLEDSNVYFWFKMRDMVLPFLDISETKNTRKEV
 LVSKDFVTLNHIRTYQAKALTYSNLSEVSRVIRVINGVTARSEVDKSLQS
 LSTMTFFLTKLAVLDKDLISKFALGKTVSQHWDEISLAFGNAPFSIKERLINRKL
 IRTENALEIRVDPVYTHDRVLSYKMSVDMPLDIRKMBETEMYNALSELVSL
 KNSDKPDVFSQMCOSLEVDPMATAKVIIVAMNSESGLTLTFEQTEANVALAODS
 EKASDGLVTPRDEEPSIKGSMARGELQAGLSDVPSESSYTRSEIESLQFHMA
 TASLIHQKWCISIVYTGPKVQOMKNFIDSLVASLSAAVSNLKLKDTAIDLETQ
 KFGVLDVASKRWLVKPSAKDHGAVVETHARKYRVALLEHDEFGIITCDNWRVAVSS
 ESVVYSDMAKRLTLRLKDGEPHVSAAKVLVDGPGCGTKTEILSRVNFEDLILV
 PGQAAMIRXRANASGIIIVATKDNVTVDSFLMNYGKRAKCOFKRFLIDEGMLHTG
 CVNPLVMSLCDIAYVYGTQOIPYINRVTFGPPYPAHFAKLEVDEVTRETLRCPAD
 VTHFLNQRYEGHVCTSSSEKSVSQEMWGAASINPVSKPLKGLITFTQSDKEALLS
 RGYADVHTVHEVQETVADVSLVRLTPTPVSIIARDSPHVLVLSLRHSKLSKTYTVVM
 DPLVSIIRDLERSVSYLLDMYKVDAGTQ"
 72. .2672
 /gene="replicase"
 /function="genomic replication, virus attenuation and
 specific protection against wild challenging viruses"
 /note="contains opal mutation"
 /codon_start=1
 /product="98.5 kDa protein"
 /protein_id="AAD44325.1"
 /db_xref="GI:5524642"

gene
 CDS
 /translation="MAYTQTATSSALLETVRGNTLVNDLAKRRLVDAADVFNARDR
 RPKVNFSEVSEQTLIATKAYPEFQITFYNTQNAVHSLAGLSLELYMMQIPYV
 SLTYDIGNFASHLFGKRAYVHCCMPNLDVDMRHEGQKDSIELYLSRERGNKHP
 NFOKEAFDAEMNEVWCHDTQTCTHSEOECTGRVYATALHSIYDIPADEFGAALL
 RKNVHCYAAHSENLLEDHVNDEINACFORQDRLTFSEFASSTLYNSHSYN
 ILKYCKTTPASNRVYKKEFLVTRVNTWFKESRIDTFLYKGVHAKGVDSQOFYK
 TMEDAHYKTKLWYNSERILLEDSNVYFWFKMRDMVLPFLDISETKNTRKEV
 LVSKDFVTLNHIRTYQAKALTYSNLSEVSRVIRVINGVTARSEVDKSLQS
 LSTMTFFLTKLAVLDKDLISKFALGKTVSQHWDEISLAFGNAPFSIKERLINRKL
 IRTENALEIRVDPVYTHDRVLSYKMSVDMPLDIRKMBETEMYNALSELVSL
 KNSDKPDVFSQMCOSLEVDPMATAKVIIVAMNSESGLTLTFEQTEANVALAODS
 EKASDGLVTPRDEEPSIKGSMARGELQAGLSDVPSESSYTRSEIESLQFHMA
 TASLIHQKWCISIVYTGPKVQOMKNFIDSLVASLSAAVSNLKLKDTAIDLETQ
 KFGVLDVASKRWLVKPSAKDHGAVVETHARKYRVALLEHDEFGIITCDNWRVAVSS
 ESVVYSDMAKRLTLRLKDGEPHVSAAKVLVDGPGCGTKTEILSRVNFEDLILV
 PGQAAMIRXRANASGIIIVATKDNVTVDSFLMNYGKRAKCOFKRFLIDEGMLHTG
 CVNPLVMSLCDIAYVYGTQOIPYINRVTFGPPYPAHFAKLEVDEVTRETLRCPAD
 VTHFLNQRYEGHVCTSSSEKSVSQEMWGAASINPVSKPLKGLITFTQSDKEALLS
 RGYADVHTVHEVQETVADVSLVRLTPTPVSIIARDSPHVLVLSLRHSKLSKTYTVVM
 DPLVSIIRDLERSVSYLLDMYKVDAGTQ"
 4906. .5634
 /gene="Mp"
 4906. .5634
 /gene="Mp"
 /function="virus movement in plants; involved in local
 lesion development and deformity of tobacco leaves"
 /note="movement protein; contains ochre mutation"
 /codon_start=1
 /product="26.8 kDa protein"
 /protein_id="AAD44328.1"
 /db_xref="GI:5524645"

gene
 CDS
 /translation="MALVVKGVKYNINEFIDLSKSEKLLPSMFTPVKSVMWKVDKIMV
 HENESLSEVNLKGVKLEGGVYVCLVGLVSGEWNLPDNCGGVSVCMVWDRKMRABE
 ATIGSYTAAAKRQFQKVPVNYGITTKDAEKNIWOVLVNIKNVMSAGYCPILSLFV
 SVCIVYKNNIKLGLREKVTSDNGGPMELSEEVDFEVMENVPMSRLAKFRKTSKSG
 PKNNNLKGGKSGRPKPKPSFDBV"
 5703. .6182
 /gene="Cp"
 5703. .6182
 /gene="Cp"
 /functions="virus assembly"
 /note="coat protein"
 /codon_start=1
 /product="17.7 kDa protein"
 /protein_id="AAD44329.1"
 /db_xref="GI:5524646"

gene
 CDS
 /translation="MSYSITSPSQFVFLSSWADPIELLNVCTNSLGNQFOQOARTT
 VQOOFSEVWKFPQSTVRFPGDVKYTRYRNAVLDPDLITALLGADFTNRRIIEVENQOS

ORIGIN		PTTAETLDATRRVDDATVAIRSAINLVNELVRGTGLYNQNTFESMSGLVWTSAPAS"	
Query Match		41.3%; Score 2621.6; DB 14; Length 6384;	
Best Local Similarity		64.2%; Pred. No. 0;	
Matches 4096; Conservative		0; Mismatches 2224; Indels 60; Gaps 8;	
QY	18	CGACAAACAAATTAACAAACAAACAAACATATTACAAACAAACAAACAAACATGGCAC	77
DB	23	CAACAAACAAACAAACAAACAAACATTTACATTTTACATTC-----TACAACTACAAATGGCAT	78
QY	78	ACATACAACTCTATAATTAGCAACGCCCTCTCTGAAGCGGTGAGTGGTGAACAACTCTCG	137
DB	79	ACACAAACAGCCACATCGTCGGCTTGTGTGAGCCGTCGAGGTAAACATACCTTGG	138
QY	138	TTAATGACCTTGAAGAGCGCATGTACGATACGCGCGTGAAGAAATTTAAGCCCGCG	197
DB	139	TCAACGATCTTGAAGAGCGCGCTATATGACGCGCGTGGATGAATTTAATGCTAGGG	198
QY	198	ACGTAGACCAAGGTCACTTTTCCAAACTATTAGCGAAGCAACACGCTTCTAGTCT	257
DB	199	ACGCGAGGCTTAAAGTCAATTTTCCAAAGTAGTAAGCGAAGAACAGACGCTTATTGCAA	258
QY	258	CCAAACGCTACCGGAGTCCAGATTACCTTTTATATATCTCAAAATGCCGTACACAGTT	317
DB	259	CCAAAGCTACCCAGAAATCCAAATTTACATTTACACACGAGAAATGCTGTGCATCCC	318
QY	318	TGGCTGGAGTTTGGAGGCAATAGAAATGGAATATCTGATGCTACAAATTTCCCTATGGAT	377
DB	319	TTGCAGGCGTCTCCGATCATTAGAAATGGAATATCTGATGATGCAAAATTTCCCTACGGAT	378
QY	378	CGCGGACATATATAGTGGGAACTTTGCGAGCACTTTGTTCAAGGAGGAGGATACG	437
DB	379	CATTGACATATGATATCGGAGGTAATTTTGCATCTCATCTGTTCAAAGGCGGAGCATACG	438
QY	438	TGCATTGCTGTATGCCAATCTGACATACAGATATAATGAGCGACGAGGACAAAGG	497
DB	439	TTCACTGCTGTATGCCAATCTGATGTCGCGACATAATGCGGACGAGGCGCAAAGG	498
QY	498	ACTCAATTTGAGATGTATTTGTCAGAAATGTCTGTTCTAAAGAGTAAATTCCTGAGTTTC	557
DB	499	ACAGTATTGAACATATACCTTTCTAGGCTCGAGAGGGGCAACAACATGTCCTCAAACTTC	558
QY	558	AAAGGAGGCTTTTAAACAGATATGCGAAGCTCCAAACGAGTCTGCTGCTCTTAAACTTT	617
DB	559	AAAAGGAAGCTTTTCCGACAGATACGCTGAAATGCGAAACGAAAGTAGTCTGTACAGATCTT	618
QY	618	TTCAGGATTTCCGAATATACCCGCGAGAGATAGTGTAGAGATACGCTGTTGCTCTGC	677
DB	619	TCCAAAGCTGTAGGCACTCTCAAGATGTTTACACGGGAGAGTGTATGCTTATTGCTTTC	678
QY	678	ACAGTTTGTATGATATTTCTGTGATGAGTTTGGAGCTGGTGTAAATCTTAAGAAATATAC	737
DB	679	ATAGTATATAGATATATCTGCGGACGAGTTCGGCGGGCACTGCTGAGAAAGAAATGTAC	738
QY	738	ATGATGTTATGACGCTTCCATTTTGGCAGAGCATTTATCTAGACCAAGGAGGTTA	797
DB	739	ATGATGTTATGCGCTTTTCCACTTTTCCGAGAAATTTACTTCTCGAAGATTTCAACGTC	798
QY	798	CGCTTAATGAATAGGCGCACTTTTCAAAGAGAGGATGATGTTCTTTTCTTTTCTTTG	857
DB	799	ACCTCGACGAGTCAATGCAATGTTTCCAAAGAGATGAGACAGGTTGACTTTTCTTTG	858
QY	858	CTGATGAAGTACTTTTAAATTTATAGTCAATAAATCAAAAATATCTTGCATTTATGTAGTTA	917
DB	859	CATCTGAGATCTCTTAATTTATAGTCAATAGTATTCTTATATTTCTTAAGTATGTTGCA	918
QY	918	AATCTTACTTCTGCTCTCTAGTAGAAATGTTTACTTTTAAAGAAATTTTATGCTACATAGG	977
DB	919	AAACTTACTTCCCAGCTCTTAATAGAGAGGTTTACATGAAGGAGTTTTATAGTAACTAGAG	978
QY	978	TTAATACTTGGTTTGTAAATTTTACAAAGTAGATACCTATATCTCTGACAGAGTTTCAAGAGTTTA	1037

DB	979	TTAATACCTGGTTTGTAAATTTCTAGAATAGATACTTTCTTATTGTACAAAGGTGTAG	1038
QY	1038	GACAACTAGGTTGTAGTAGTGCATCTTATGAGCGGATGGAAGAGCGCTTTGCTTACA	1097
DB	1039	CGCATAAAGGTTGTAGTAGTGCATCTTTTCAAGAGTATGGAAGAGCGATGCGACTACA	1098
QY	1098	AGAAACCTTTGGCATGTTTCAACACATGAAGAGCAATCTTTAGAGACACCGCTTCGGTTA	1157
DB	1099	AAAGACTCTTGGCATGTTTCAACAGTGAAGAAATCTTGTAGAGGATCTTCAATCAGTTA	1158
QY	1158	ACTTTTGGTTTCCCTAAAGATGAAGGACATGGTGTAGTACCGCTGTTTGGAGGTTCTATTA	1217
DB	1159	ATTACTGGTTTCCAAAATGAGGATATGGTGTAGTTCACATATTGACATATCTCTCG	1218
QY	1218	CMAGCAAAAGATGACAAAGAGTGAAGTCAATGTTTAATCGTGAATCTGCTTTTACAGATGC	1277
DB	1219	AGACTAATAAAGAAACACGCAAGAGGCTTCTAGTTTCAAAGGACTTTGTTTATACAGTGT	1278
QY	1278	TTAATCATATCAGAACATATCAAGCCAAAGCGTTAACTTTACCAAGAACGTTATTTCTTCG	1337
DB	1279	TAAATCACATTTCTGATCCAGGCCAAAGCGCTTACTTACTCCAGCTGTATCTTTTCG	1338
QY	1338	TGAGTCTTATAAGATCCCGGTGATTAATCAATGGTGTACTGTAGTCTTGAATGGGATG	1397
DB	1339	TCGAATCAATTCGTTGAGAGTGCATCAATTAACGGGGTTACTGTAGTCTGAGTGGGATG	1398
QY	1398	TAGTAAAGCAATCTTCAACCTTGTCAATGACTTTCTTCTTGGCAGACTAAGCTGGCTG	1457
DB	1399	TCGATAAATCATTTATTTACAGTCTTCTGATGACGTTCTTCTTACACACCAACGCTTGGCG	1458
QY	1458	CGCTTCAAGACGATATAGTAATGGGAAGTTTTCGGTCTTGGATAAGACCACCTTCTGAAC	1517
DB	1459	TTCTGAAGACGATCTTTTGAITAGCAAGTTTGCACTTGGACCAAAAACCTGTCTCAACAC	1518
QY	1518	TTATTTGGGATGAGGTGGGCAAAATTTTGGAAACGTTTTCGCCACTATCAAAAGAGAGAT	1577
DB	1519	ATGTGTGGATGAGATTTCCCTAGCTTTCGGCAATGCTTTCCCATCGATCAAGAAAGAT	1578
QY	1578	TGTTGAGCAGGAAATTTCTGATGTAAAGTGAAGTGTCTGAAGATCAAGATCCAGATC	1637
DB	1579	TGATAAACCGGAACTGATCAGAAATTTACGGAAGTTCGTTAGAGATCAGGCTGCCGATC	1638
QY	1638	TGTTATCTCATGGAAGACAGGTTCTGATGCTCAATACCAAGTCTGAGGAGTTACCGC	1697
DB	1639	TTTATGTCTTCTTCCATGATGAGTTAGTTTCTGAGTACAAAATGTCTAGTGGACATGCCG	1698
QY	1698	ATCTAGATATCAAGAGGACTTAGAAGAGCTGAGCAAAATGTACGCGCTTATCAGAAAT	1757
DB	1699	TGCTAGACATTAGGAAAGATGGAAGAACTGAGGAAATGTACAAATGCACTGTCCGAAC	1758
QY	1758	TATCTATCCTTAAGGTTGTGATTAATTTGATATCCGGAAGTTCAAGAGATGTGCAAGG	1817
DB	1759	TGTTCTACTTAAAAATTCAGACAGTTCGATGTGACGTTTTCCTCCAGATGTGCCAAT	1818
QY	1818	CTTTAGATGTTAGTCTCTGATGTGGCAGCAGGTAATCGTTGAGTGGCCGAGAAATAGAA	1877
DB	1819	CTTTAGAGTTCGATCCAAATGACTGACGCAAAAGGTAATAGTAGCAGTTATGAGCAACGAG	1878
QY	1878	GCGTTTAACTCTTACTTTTGAAGCAACCCGAGGAGAAATGTGGCTTAAGGCTCTTAAA-	1936
DB	1879	GTGGTCTTACTCTCACTTTGAACGCCCAACCGAGCTAATGTTCGCTAGCTAGCTGCAAG	1938
QY	1937	-----AGCACGGCTCTGAGGCCGCTGTATGTCTTGAACCGCATCCGAGAGAGGTGAACG	1991
DB	1939	ATTCTGAAAGGCTTCTGATGGGGCTTGGTAGTTACCCCAAGAGATGTTTGAAGAACCGT	1998
QY	1992	TAAATAAATTTCTTATTTGAGAAAGGAGATTTGCTGTGTGCGAAGAAAGTCTATGGTT	2051
DB	1999	CCATNAAGGTTCCGATGGGCCGCTGGTGAAGTTCAATTTGGCCGGAATATTCTGGGACGCTC	2058
QY	2052	TGACGAATGCTAACTTTAGAGCACGAGGTTGGAGTCCCTCAACGATTTCCATAAGGCTT	2111
DB	2059	CTGAATCTTCACTACACTAGGAGCGAGGAGATTGAGTCTCTCGAGCAGTTTCAATATGCAA	2118

Qy 2112 GCGTGGATAGTGTGATTAACAAGCAATGGCATCGGTGTCTACACTGGCTCACTCAAAAG 2171
Db |||||
Qy 2119 CAGCTAGTTCGTAAATTCATAAGCAGATGTGTCGATCGTGTACAAGGCCCTCTTAAAG 2178
Db |||||
Qy 2172 TTCAACAATGAAGAACTATGTGACACATTTGGCAGCTTGGTGTCCGCCACTGTATCAA 2231
Db |||||
Qy 2179 TTCAACAATGAAGAACTTTATAGACAGCCCTGGTAGCTCGCTCTCTGCTCGCGGTGCGA 2238
Db |||||
Qy 2232 ATCTATGCAAGTCACTAAAGATGAAGTCCGGTATGATTTCTGATTCAGGAGAGAAAGTGG 2291
Db |||||
Qy 2239 ATCTAGTGAAGATCCTTAAGATACACCGCGATTTGACCTTGAACCTCGTCAPAAGTTGG 2298
Db |||||
Qy 2292 GTGTTTGGGATGTCATTTTGAAAGTGGTCTCTCAAACTCGCGGCCCAAGGTCATTCAT 2351
Db |||||
Qy 2299 GAGTTCCTGGATGTGCTTCGAAAGTGGCTAGTTAAACCATCGCAAGACCATGCAT 2358
Db |||||
Qy 2352 GGGGAGTTGCTCGATTAACAGGGGAAATGTTTACTGCATCTCTATCTATTAAGAGGAG 2411
Db |||||
Qy 2359 GGGGGGTTGTGAGACTCATCGGAGAAATATACGCGTCGCATTAATCTGGAGCACGATGAAT 2418
Db |||||
Qy 2412 ATAGAATGGTGACTGAGAGGACTGGAGAGGGTGGCTGTATCATCTGTATACATATGAT 2471
Db |||||
Qy 2419 TTGSCATTTATCGTGCATTAATCTGGCGACGGGTGGCTGTGAGTTCTGAGTGGTAGTAT 2478
Db |||||
Qy 2472 ATTCTGATATTGCAAGCTCCAAATCTGAGGAAACAAATGAGAGACGGTGAACCCCAAG 2531
Db |||||
Qy 2479 ATTCTGATATGGCTAAACTCAGGACTCTGAGAAGATTGGCTCAAAAGATGGAGAACCAACAG 2538
Db |||||
Qy 2532 AACCTACTGCAAGATGGTACTTGTGATGGGTGGCTGTGTTGTTGGAAGTACAAAGGAG 2591
Db |||||
Qy 2539 TTAGTTTCAAGAAAGTGGTTTGTGATGGCTTCAGGGTTCGCGGAACAAAGGAAA 2598
Db |||||
Qy 2592 ATTTTGAAGATTTGATCTTGAAGGATTTGATCTTGGTTCTTGGAAAAACAAGTCTGCTG 2651
Db |||||
Qy 2599 TTCTTTGAGNGTTAAATTCGAAGAACCTTAATCTTGTCCTCGTGGCTCAAGCTGCCG 2658
Db |||||
Qy 2652 CTATGATCAGAAGAGGGCTAAATTCATCTGACTGATTAAGAGCCACAAATGACAAATGTA 2711
Db |||||
Qy 2659 AGATGATCAGATGAAGAGCTAAATCGTGGCGCATTAATAGTGGCTTACAAAGGATAATGTC 2718
Db |||||
Qy 2712 GAAGGTAGATTCACTTCTAATGATCCAAACCGCGATCAC-----ACAAAGGC 2762
Db |||||
Qy 2719 GCACCGTCGATTCATTTTGTATGAATTAACGGGAAAGGGCACGCTGTCAGTTCAAAAGAT 2778
Db |||||
Qy 2763 TTTTATTTGATGAAGGTGTGATCTGCACACCGGTTGTGTAACTTCCCTGGTGTCTTATCT 2822
Db |||||
Qy 2779 TGTTTCAAGACGAGGTTTGTGCTGCTGCTACTGCTGTGCTGCTGCTGCTGCTGCTGCTGCT 2838
Db |||||
Qy 2823 CTGTTTGGGACATCGCATATATTTACGGAGATACACAGCAGATTCCTTTTCAATTAACAGAG 2882
Db |||||
Qy 2839 CTCTGTGCGATATTGCAATATGTTTATGGAGACACCCAAACAGATTCCTGATACATCAACAGAG 2898
Db |||||
Qy 2883 TTTCAAGATTTCCGTATCCCAACATTTTGGAGAGCTGCAAGTGGATGAAGTTGAGATGA 2942
Db |||||
Qy 2899 TAACGTGTTTCCCGTACCTTCGACACTTTGCAAAATTTGGAGGTGCGAGCAGTGGAAACAA 2958
Db |||||
Qy 2943 GGAGGACACACTGAGATGCCAGGTGATGTAATTTTTTCTTACAACTCGAAGTACGAAG 3002
Db |||||
Qy 2959 GAAGAACTACTCTCTGTTGTCGGCTGATGTCACACACTTCTTAAATCAAAAGGTATGAAG 3018
Db |||||
Qy 3003 GAGGGGTGACAAACACTTCAACTGTACAACGATCGGTCTCATCTGAGATGATAGGGGGTA 3062
Db |||||
Qy 3019 GACACGTAATGTGACGCTCTCTGAAAGAAATCAGTTTCCAGGAAATGGTTAGTGGGG 3078
Db |||||
Qy 3063 AGGAGTACTAAACAGAGTTTCCAAACCACTTAAAGGGAAATTTGTAACCTTCACTCAGG 3122
Db |||||
Qy 3079 CTGCGTCTATCAATCTCTGTGTCGAAGCCGCTTAAAGGAAATTTTGTACTTTACACAGT 3138
Db |||||
Qy 3123 CTGATAAATTTGAGTGTAGAGGAGAGGGCTATTAAGAAATGTGAACACCGGTTCACTGAGATCC 3182
Db |||||
Qy 3139 CTGACAAGGAGGCCCTTCTCTCAAGGGGCTACCGCAGATGTCCTACTGTACATGAGGTAC 3198
Db |||||

Qy 3183 AAGGAGAAACCTTTTGAAGATGTGTCGTGGTCAAGTTGACGGCAACTCCACTGACTCTGA 3242
Db |||||
Qy 3199 AAGGTGAGACTTATGACAGAGCTATCGTTAGTTCCGATTAACACACTACGCTGTATCTATCA 3258
Db |||||
Qy 3243 TTTTCCAAAGTCTTCCCGCATGTTCTAGTCCGCTCTGACTAGACACACAAAGAGCTTCAAAAT 3302
Db |||||
Qy 3259 TCCGAAGAGACAGTCCGCGATGTTCTGGTCTCGTTGTCAAGACACACAAAAATCCCTAAAGT 3318
Db |||||
Qy 3303 ATTACACCGTAGTGTGTAGATCCCTTTAGTACAGATAAATTAAGTATGTTCTTCTTTTAAGCT 3362
Db |||||
Qy 3319 ACTACACCGTTGTGATGGATCCCTTTAGTTAGTATCATTAGAGATTTAGAACGGGTTAGTA 3378
Db |||||
Qy 3363 CTTTCTTTTGAAGATGTATATGTTAGAGCAGGTAGTAGATAGCAATTAACAGATGATG 3422
Db |||||
Qy 3379 GTTACTTATTAGACATGTCACAAAGTAGTGCAGGTACTCAATAGCAATTAACAGTGCAGCT 3438
Db |||||
Qy 3423 CAGTGTTCAAAGTTCATAATCTCTTTGTGCAACACACTAAATCAGAGAGCTTTCCAGATC 3482
Db |||||
Qy 3439 CTGTGTTTAAAAATTTCAATCTTTTGTAGCAGCTCCAAAGACTGGAGATATATCTGATA 3498
Db |||||
Qy 3483 TACAGTTCATTAACGATGTATGCTCCCTGGTAAATAGTACTATCTTAAACAAGTATGATG 3542
Db |||||
Qy 3499 TGCAAATTTTACTATGATAGTGTCTTCTGGGAACAGACGTTGTGAAACAACATACGACG 3558
Db |||||
Qy 3543 CTGTTACCATGAGGTTACGTTGATAATAGTCTTAAATGTGAAGGATGTGTTCTTTGATTTTT 3602
Db |||||
Qy 3559 CTGTTACCATGAATTTGACTGACATTTCTGAAATGTCAAAAGATTGCAATATTAGATATGT 3618
Db |||||
Qy 3603 CCAAAAGTATTCGATGCGCAAGGAGGTGAAACCATGCTTAGAGCCAGTTTTCGCTACCG 3662
Db |||||
Qy 3619 CTAAGTCTGTAGTGTCTCCGAAAGATGTCAAAACCAACTTTAATACCAGTGGTACGAACCG 3678
Db |||||
Qy 3663 CGCGCAACCGCCCAAGGGCTGCAGGACTACTCCAAAATCTGGTTGCAATGATTAAGAAAGAA 3722
Db |||||
Qy 3679 CGCGCAAAATGCTCCCGACACTGGACTGTGGAATCTAGTTGCGATTAAGAAAGAA 3738
Db |||||
Qy 3723 ATTTCAACGACACAGACCTGACCGGGACGATTGACATTTGAGAGACCGCATCTGTTGTAG 3782
Db |||||
Qy 3739 ATTTTAAATTCACAGAGTTGTCCGAGTAGTTGATATTGAAATACTGCACTCTTTAGTGG 3798
Db |||||
Qy 3783 TAGATAAGTTTTTTGATAGCTATTTTAAAGAAAGAAATAACACAAATAATTTGCTG 3842
Db |||||
Qy 3799 TAGATAAGTTTTTTGATAGTTATTTTAAAGAAAGAAAGAAACCAACCAAAATTTTTT 3858
Db |||||
Qy 3843 GAGTGTAGCAGAGGATTCATATGATGAGATGTTGGAAGACAGGAAGAGTACTATTGG 3902
Db |||||
Qy 3859 CACTGTTTAGTAGAGAGTCTCTCAATAGGTGGATAGCAAAAGCAAGAACCAATTCG 3918
Db |||||
Qy 3903 ACGACTTGGCTAATCAATTTTACAGATCTGCCGCCATCGATCAGTACAGACACATGA 3962
Db |||||
Qy 3919 GTCAGATGGCCGATTTTGTATTTTGTGGATCTTCCAGCCGTTGATCAGTACAGGCATATGA 3978
Db |||||
Qy 3963 TCAAGGCTCAACCAAAACAGAAATTCGACCTTTCAATTCAGATGAATACCCGTCTGTCG 4022
Db |||||
Qy 3979 TTAAGCGCAACCGAAGCAGAAACTGGATCTGTCAATTCAGACAGAAATTCAGAGGTTCG 4038
Db |||||
Qy 4023 AAAACAATTTGTCTACCAATTCGAAGCAGATCAACGGTATTTTGG---COGGTTTCTCAGAGC 4079
Db |||||
Qy 4039 AAACGATTTGTATCATCTCAAGAAATCAACGCAATATTGGTCTCTTTTCAGTGAGC 4098
Db |||||
Qy 4080 TTAACAGGTTGTGCTCGAGGCATTTGATTTCTAAGAAAGTTTCTTTTCTTTTACTAGGAAA 4139
Db |||||
Qy 4099 TTAACAGGCAATTTACTTGAAGTATTGACTCAAGCAGATTTCTTTGTTCTTTTACGAGAAAGA 4158
Db |||||
Qy 4140 CTCGCAACAGATTTCAAGAAATTTTCTCGATCTCGACTCGCAGCTTCTTATCGATGTGT 4199
Db |||||
Qy 4159 CACCGCTCAGATCGAAGATTTCTACGGAGATCTAGCAATCATGTGCCCAATGGACGTAC 4218
Db |||||
Qy 4200 TAGAACTGGATTTCTTAAAGTATGATTAAGTCAACAGAACGAGTTTCAATTTGTGCTGTAGAGT 4259
Db |||||
Qy 4219 TTGAGTTGGATGTTTCAAGATGATGAAGTCTCAAAACGAGTTTCAATTTGTGCTGTGAGT 4278
Db |||||
Qy 4260 ATGAATAATGGAAGAAAGATTGGGTCTCAATGAGTTTTTGGCCGAGGTGTGGAAACAAGGC 4319
Db |||||

Db 4279 ACAGAACTCGAGAGACTGGGTCTGGAGGATTTCTTGGCAGAGTGTGGAAACAAGGGC 4338
QY 4320 ACAGGAAAAACAATTTGAAGGATTACATTTCTGGAATCAAGACATGTCTGTGGTATCAAA 4379
Db 4339 ATAGAAAAACCACTCTGAAAGATTACATCTGCTGTATAAAAACGTGTATTGTGTACAGA 4398
QY 4380 GGAAGAGCGTGTATGTACTCTTTTCATCGGCAATCTGTATTAATAGACGCTGTCTGG 4439
Db 4399 GAAAGAGTGTGTATGTATCAAACTTTTATCGGTAATACCGTCATCTGCTTCTGTCTAG 4458
QY 4440 GTTCAATGTTACCGATGGAAGAGTCAATAAGGTGCTTTTGTGGAGAGGATTCGGTTT 4499
Db 4459 CATCAATGCTCCGATGGAATAATTTGATAAAGAGCCTTCTCGGAGATGACAGTTGT 4518
QY 4500 TGTATTTTCCAAAGGTTTGGATTTCCCTGACATTCAGTCAATGCTTAATCTCATGTGGA 4559
Db 4519 TGTACTTTCTAAGGTTGTGAGTATCCGATATACAAAGCTGCTAATCTAATGTGGA 4578
QY 4560 ATTTGAGGCCAACTGTATAGAAAGAGGTACGGTTACTTTTGTGTGTAGATACATCATAC 4619
Db 4579 ATTTGAGGCCAACTGTTCAAGAACCAATATGGGTACTTCTCGGGAGGTACGTGATTC 4638
QY 4620 ACCATGATAGGAGCAATAGTGTATATCATCTTTGAGTTGATCTCCAACTTGGGG 4679
Db 4639 ATCAGATAGAGTTGCATAGTATATACACCTTTTGAAGCTGATTCGAAACTTGGTG 4698
QY 4680 CAAAAACATATCAAGGATTTATGATCACTTAGAAGGTTAAGGTGTCTTTGTGCGATGTTG 4739
Db 4699 CTAAACACATCAAGGATTTGGATCACTTTGAGAGGTTTCAAGATCCCTCTGTGTGTTG 4758
QY 4740 CTTGTTGCTCGAAACTGTGTAGGCTTTCCGAGCTGGAACGAGCTATCAAGAGG 4799
Db 4759 CTGAGTCTGT--GAACAAATGCGCGTATTAACACAAATTTGACGACGCTGTGGGAGG 4815
QY 4800 TTCAATAACCGGATGTAGTGTCTGTCTTTTAAATTTGATGTTAACAATAATTTCTGTG 4859
Db 4816 TTCAATAACCGGCCACCTGTTGTTTGTATAGAGTTAGTTAAGTATTTGTGTAG 4875
QY 4860 ATAAATTTTATTTAGAACTTTGTTTAAATGGCTGTAGTCTCAGAGATATCTGCAAA 4919
Db 4876 ATAAAGTTTGTATTAGAGTTTATTTCTGATGGCTCTAGTGTAAAGGTAGATTAAT 4935
QY 4920 ATTAGCGAGTTCATGTATCTTTGAAACAGGATGAGATATCTCCGCAATTCATGACTAAG 4979
Db 4936 ATTAATGAGTTTATCGATCTGTCAAAGTCTGAGAACTTCTCCGCTCGATGTTCAAGCCT 4995
QY 4980 GTCAAGAGTGTAGAAATATCGACTGTGGACAAGATTATGCTGTTAAGATGATAGTCTT 5039
Db 4996 GTAAAGAGTGTATGTTTCAAAAGGTTGATTAAGATTAATGCTCCATGAAATGAATCATG 5055
QY 5040 TCTGATGTAGATTTACTTAAAGGTTTAAAGTTAGTTAGTTAAGAAAGGATATGTTGCTAGCT 5099
Db 5056 TCTGAAGTAAATCTCTTAAAGGTTGATAAATCTTATAGAGGTGGTATGTTTCTTAGTC 5115
QY 5100 GATTTGCTAGTGTCTGGGAGTGGAAATCTCCCGATAACTGCCGTGGTGGTGTGAGTGT 5159
Db 5116 GGTCTTGTGTCTGCGGTGAGTGGAAATTTACCAGATAATTTGCGGTGGTGGTGTGAGTGT 5175
QY 5160 TGTATTTAGATTAAGAAATGAAGAGTAAAGAGTAAAGCAACGCTGGTGGGTATACAGCC 5219
Db 5176 TGCATGTTTGAAGAAGATGGAAGAGCGGACGGAAGCCACACTCGGGGTATATTAACAT 5235
QY 5220 CTTGCTTGCAAAAAAGAAATTTCTTTTAAAGCTAATCCCTAATTTTCAATAAATCCGAG 5279
Db 5236 GCTGCTGCTAAAGACGGTTTCAGTTTAAAGTGTCCCAATTTACGGTATTTACAAAG 5295
QY 5280 GATGCTGAGAGCAACCGTGGCAGTGTATGTAATCAAGAGAGTGGCTATGGAAGAA 5339
Db 5296 GATGCAAGAAAGAACATATGCGAGTCTTAGTAAATATTAATAATGTAATAATGAGTGC 5355
QY 5340 GGATACGTCTTTATCTTTGGAGTGTCTTTTCAATTTGTGTATGATACATAAAAAATATGTA 5399

RESULT 11
AF332868

Db 5356 GGCTACTGCCCTTTGTTCATTAGAAATTTGTGTGTGTATTTGTTTATAAATAATATA 5415
QY 5400 AGAAAAAGTTTGAAGGAAAGTATTTTGTAGTGTGACAGCGGCTCCGCAATTTGAACCTCACT 5459
Db 5416 AAATTTGGTCTTGAAGGAGAAAGTAAACGAGTGTGAACGATGGAGGACCCATGGAATTCG 5475
QY 5460 GAAAAGGTTTTCAGGAGTTCGTGGATGAAGTACCAATCGCTGTGGAACCTCGAAAAGTTT 5519
Db 5476 GAAGAAGTTTGTGATGAGTTCATGGAGAAATGTTTCCAATGTCCGTTAGACTCGCAAGTTT 5535
QY 5520 CCGGAAAAAAGAAATGGTAGTAATAATGTTTAAATAA 5561
Db 5536 CGAACCAATCTCAAAAAGAGGTCCGAAATAATAATAATTTAGTAGGGGCGGTTCA 5595
QY 5562 -----AAGAAAAATAATAACAGTGTGAAGGGTTTTAAATTTAGGAAAAATGAGGAT 5615
Db 5596 GCGGAGAGCCATAAACCAAAAAGTTTGTAGTAAATTTAAAGAGTTTAAATAATTTGATT 5655
QY 5616 AATGTAAGTGTAGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATCCCTTATA 5675
Db 5656 AAGATTAAAGCCGAGACGTCGGTCCGGAATCTGATTCGTAATTA 5712
QY 5676 CAATCAACTCTCCGAGCCAAATTTGTTTACTTATCTTCCGCTTACGAGATCCCTGTGAGC 5735
Db 5713 CAATCACTCTCCATCGCAATTTGTTTGTCTGTATGGGCTGACCTATAGAAT 5772
QY 5736 TGATCAATCTGTGTACAAATGCAATTTGGGTAACCAAGTTTCAACGCAACAGCTAGAGCAA 5795
Db 5773 TGTAAACGTTTGTACAAATTCGTTAGGTAAACCAAGTTTCAACACAGCAAGCAAGAACTA 5832
QY 5796 CAGTCAACAGCAATTTGGGATGCTGGAACCTGTGCTAGTATGACAGTACAGATTTTC 5855
Db 5833 CTGTTCAACAGCAGATTTCAAGAGGTGTGGAACCTTTCCCTCAGAGCAACCGTCAGATTTTC 5892
QY 5856 CTGATCGGATTTCTATGTGTATAGATATAATTCGACGCTTGAATCGTTGATCAACGCGCT 5915
Db 5893 CTGGCAATGTTTATAGGTGTACAGGTACATGCAAGTTTGTAGATCTCTAATTTACTGCGT 5952
QY 5916 TATTAATAGCTTTGTATCTAGAAATAGAAATATAGAGTTTGAATATCAACCCGCAACCGA 5975
Db 5953 TGCTGGGGCTTTTGATACTAGGAATAGAAATATCGAAGTAGAATAACACAGCAGAGTCCGA 6012
QY 5976 ATACTACTGAATCTGTAAACGACTCAGAGGTAGACGATGCTACTGTAGCTATAAGGG 6035
Db 6013 CAACAGCTGAAACGTTTAGATGCTACCCGAGGGTAGACGCTACGGTTGCAATTCGGT 6072
QY 6036 CTTCAATCAATATTTGGCTTAATGAACCTGTTCTGCGAACTGGCATGTTCAATCAAGCAG 6095
Db 6073 CTGCTATAAATTAATTTAGTTTATGAACCTAGTAGAGGTACTGGACTGTACAAATCAGATA 6132
QY 6096 GCTTTGAGACTGCTAGTGGACTTGTCTGGACCAAACTCCGGCTACTTTAGCTATTGTTGT 6155
Db 6133 CTTTGAAGTATGCTCGGTTGGTCTGGACCTCTGCACCTGCATCTTAAATCATAGGT 6192
QY 6156 GAGATTTCTTAAATGAAGTCTGGAAGTCTTAAATTTCAAGGTGGTGTGATACCAAAATC 6215
Db 6193 GCTGAAATATAA-----GTTTGTGTTTCTTAAACACACACGCTGTAACGATGAACG 6244
QY 6216 AGCAGTGTGTTGTTCCACTTAAATATAACGATTTGTCATATCTGGATCCAAAGTTTAAA 6275
Db 6245 TACAGTGTGTTTCCCTCCACTTAAATCGAAGGGTGTGTTCTTGAGCGCGCGAGTAAAC 6304
QY 6276 CCATGTGATGTTGTATATCTGTGTATGCGGTGTAAGCATCGGAGAGTTTGAATCTCTCCC 6335
Db 6305 ATATATGTTTTCATATATGTCGAGTACGCACTGTAAGGAGGATTCGAATTCCTCCCG 6364
QY 6336 TAAACCCCGGTAGCGCCCA 6355
Db 6365 GAACCCCGGTTGGGGCCCA 6384


```

LOCUS      AF332868                6383 bp ss-RNA      linear      VRL 12-MAR-2001
DEFINITION Tomato mosaic virus isolate Queensland, complete genome.
ACCESSION  AF332868
VERSION    AF332868.1  GI:13274576
KEYWORDS   Tomato mosaic virus
SOURCE     Tomato mosaic virus
ORGANISM   Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE  1 (bases 1 to 6383)
AUTHORS   Lee, S.M., Smith, G.A. and Polkinghorne, I.G.
TITLE     Complete nucleotide sequence of tomato mosaic virus Australia isolate
JOURNAL
REFERENCE  2 (bases 1 to 6383)
AUTHORS   Lee, S.M., Smith, G.A. and Polkinghorne, I.G.
TITLE     Direct Submission
JOURNAL   Submitted (29-DEC-2000) Department of Biochemistry, The University of Queensland, QABC, Gehrman Labs, The University of Queensland, St. Lucia, Brisbane, QLD 4072, Australia
FEATURES   Location/Qualifiers
            1..6383
             /organism="Tomato mosaic virus"
             /viral
             /mol_type="genomic RNA"
             /isolate="Queensland"
             /specific_host="Licopersicon esculentum cv. Grosse Lisse"
             /db_xref="taxon:12253"
             /country="Australia; Queensland"
            72..4922
             /note="involved in genomic replication; contains readthrough stop codon"
             /codon_start=1
             /transl_except=(pos:3420..3422,aa:OTHER)
             /product="184 kDa replicase"
             /protein_id="AAK17988.1"
             /db_xref="GI:13274578"
             /translation="MAYTQTATSSALLETVRGNNTLVNDLAKRRLDYTVADFNFARDR
RPKNVFSKVSSEETLIATKAYPEFOITFYNTQNAVHSLAGLSLELYLMQIPIYG
SLTVDIGNFASHLFGKRAYVHCMPNLDVDRIMRHEGQKDSIELYLSRLERGNKHP
NFKQEAADFAEMPNVCHDTQTCRHSOECYTGRTVVAALHSIVDIPADEFGAALL
RKNVHCYAFNSRENILLDESDSHVNLDEINACFORQDGLTTFSESTLNSHSYVN
ILKYCKTYFPASNSREVMKELFVTRVNTWFKPSRIDTFLLYKGVHAKGVDSQFTK
AMEDAWHYKKTILAMCNSERILLESSSVNVWPKMRMDVIVPLFIDISLETSKRTRKEV
LVSKDQFVDSQMSQLEVDVPMIAKVIIVAVMSNESGLITLTFEQPTAEANVALAODS
TASLIHKQMCISIVYTGPKVQOMKQNFIDSLVASLSAASNVLKILKOTAAIDLETRQ
KFGVLDVASKRWLVKPSAKNHGAVVTHARKVHVALLEHDEFIITCDNRWRVAVSS
ESVYSDMAKRLTLRLDKGEPHVSAAKVLVDGPGCGTKKEILSRVNFEDLIIV
PGQAAEMIRRRANASGIIVATKDNVTVDLSFLMNYGKGARCOFKRLFIDEGMLHTG
CVNPLVMSLCDIAIYVGDYQQIPIYINRVTFGPPYPAHFAKLEVDETRRTILPCPAD
VTHPLNQRVGHVQVETSSSKSVSQEMVGAASINPVSKPLKGLITFTQSDKEALLS
RGYADVHTVHEVQGVETADVSLVRLTPTVSIITARDSPHVLVSLRHTKSLKYTVVM
DPLVSIIRDLERSVGLVRLTPTVSIITARDSPHVLVSLRHTKSLKYTVVM
YKCLPGNSLTLLNNDYVTKMLTDIISLVKDCILDMSKSVAAKPVKPTLLIPWRTAA
EMERQTLLENLWAMIKRNPNSPELSGVDIENIASLVDPKFDPSLILKEKRNKNF
SLFRESLNIWAKQVQITGLADFDVPLPAVDQTRHMKIAQPKQKLDLSIQTEYP
ALQITVYHSHKINAIFGLSELTQLDLSIDSRFLFFTRKTPAQIETDFDGLDSHV
PMDVLEDVSKYQNSQNEPFCAYEIRWRLGLEDFLAEVKQGHRTKTLKDYTAGIK
TCLWYQRKSGDVTFFIGNVTIIASCLASMLPMEKLIKGAFCDDSLIYFPKGCYEDH
QOANLWMTPEAKLFGKQYGVFGFCYVIHHRDGCIVYDPLKILSKLGAKHIKWDHIL
EERPSILCDVAESLNCAYITQLDDAVGEVHKTPAGPSFVYKSLVKYLSKVLFRSLF
LDGSSC"
            72..3422
             /note="involved in genomic replication"
             /codon_start=1
             /product="126 kDa replicase"
             /protein_id="AAK17988.1"
             /db_xref="GI:13274577"
             /translation="MAYTQTATSSALLETVRGNNTLVNDLAKRRLDYTVADFNFARDR
RPKNVFSKVSSEETLIATKAYPEFOITFYNTQNAVHSLAGLSLELYLMQIPIYG
SLTVDIGNFASHLFGKRAYVHCMPNLDVDRIMRHEGQKDSIELYLSRLERGNKHP
NFKQEAADFAEMPNVCHDTQTCRHSOECYTGRTVVAALHSIVDIPADEFGAALL
RKNVHCYAFNSRENILLDESDSHVNLDEINACFORQDGLTTFSESTLNSHSYVN
ILKYCKTYFPASNSREVMKELFVTRVNTWFKPSRIDTFLLYKGVHAKGVDSQFTK
AMEDAWHYKKTILAMCNSERILLESSSVNVWPKMRMDVIVPLFIDISLETSKRTRKEV
LVSKDQFVDSQMSQLEVDVPMIAKVIIVAVMSNESGLITLTFEQPTAEANVALAODS
TASLIHKQMCISIVYTGPKVQOMKQNFIDSLVASLSAASNVLKILKOTAAIDLETRQ
KFGVLDVASKRWLVKPSAKNHGAVVTHARKVHVALLEHDEFIITCDNRWRVAVSS
ESVYSDMAKRLTLRLDKGEPHVSAAKVLVDGPGCGTKKEILSRVNFEDLIIV
PGQAAEMIRRRANASGIIVATKDNVTVDLSFLMNYGKGARCOFKRLFIDEGMLHTG
CVNPLVMSLCDIAIYVGDYQQIPIYINRVTFGPPYPAHFAKLEVDETRRTILPCPAD
VTHPLNQRVGHVQVETSSSKSVSQEMVGAASINPVSKPLKGLITFTQSDKEALLS
RGYADVHTVHEVQGVETADVSLVRLTPTVSIITARDSPHVLVSLRHTKSLKYTVVM
DPLVSIIRDLERSVGLVRLTPTVSIITARDSPHVLVSLRHTKSLKYTVVM
YKCLPGNSLTLLNNDYVTKMLTDIISLVKDCILDMSKSVAAKPVKPTLLIPWRTAA
EMERQTLLENLWAMIKRNPNSPELSGVDIENIASLVDPKFDPSLILKEKRNKNF
SLFRESLNIWAKQVQITGLADFDVPLPAVDQTRHMKIAQPKQKLDLSIQTEYP
ALQITVYHSHKINAIFGLSELTQLDLSIDSRFLFFTRKTPAQIETDFDGLDSHV
PMDVLEDVSKYQNSQNEPFCAYEIRWRLGLEDFLAEVKQGHRTKTLKDYTAGIK
TCLWYQRKSGDVTFFIGNVTIIASCLASMLPMEKLIKGAFCDDSLIYFPKGCYEDH
QOANLWMTPEAKLFGKQYGVFGFCYVIHHRDGCIVYDPLKILSKLGAKHIKWDHIL
EERPSILCDVAESLNCAYITQLDDAVGEVHKTPAGPSFVYKSLVKYLSKVLFRSLF
LDGSSC"
CDS
Query Match      41.2%; Score 2620.6; DB 14; Length 6383;
Best Local Similarity 64.2%; Pred. No. 0;
Matches 4093; Conservative 0; Mismatches 2229; Indels 55; Gaps 8;
QY 18 CGACACAAACAATAAACAACAAACAATATACAAACAACAAACAACAAACAATGCAC 77
DB 23 CAACAACAACAACAACAACAATATACATTTTACATTC---TACAACATACAATGCAT 78
QY 78 ACATACAATCTATAAATAGCAACGCCCTTCTTGAAGCGTGAGTGGTAAAAACAACCTCG 137
DB 79 ACACAACAACGACCAACATCGTCGCTTGTGAGACCGTCCGAGGTAAACAATACCTGG 138
QY 138 TTAATGACCTTGAAGAAGCGCATGTACGATACGCGCGTGGAGAAGATTTAAACCCCGCG 197
DB 139 TCAACGATCTTGAAGAAGCGCGCTTATATGACACAGCGGTAGATGAATTTAATCCTAGG 198
QY 198 ACCGTAGACCAAGGTCAACTTTTCCAAACTATTAGCGAAGACCAACGCTTCTAGTCT 257
DB 199 ACCGAGCGCTTAAAGTCAATTTTCCAAAGTAGTAAGCGAAGAACACAGCGCTTATGCAA 258
QY 258 CCAACGCGTACCCGGAGTCCAGATTACCTTTTATATACTCTCAAAATCCCGTACACAGTT 317
DB 259 CCAAGCTTACCAGATTCCAAATTACATTCTACAACACGACAGATGCTGTGATTCCTCC 318
QY 318 TGGCTGAGGTTTGAGAGCATTAGAATTTGGAATATCTGATGCTACAAGTTCCCTTATGAT 377
DB 319 TTGAGCGGCTCTCCGATCATTAGAATTTGGAATATCTGATGATCAAAATTCCTTACCGAT 378
QY 378 CGCGGACATATATAGTGGGAACCTTTGACGACACATTTGTTTCAAAGGACGAGGATTACG 437
DB 379 CATTGACATATGATATCGGAGGTAAATTTTGCATCTCATCTGTTCAAAGGGCGACATACG 438

```


QY 438 TGCATTCCTGTATGCCCAATCTGGACATACGAGATATAATGAGGACGAAAGGCAAAAGG 497
DB 439 TTCACCTGCTGTATGCCCAATCTAGATGTCGCGACATAAATGCGGACACGAGGCGCAAAAGG 498
QY 498 ACTCAATTTGAGATGTATTTGTCAGATTTGTCGTTCTTAACAAGGTAATTTCTGAGTTTC 557
DB 499 ACAGTATTGAACATATACCTTTCTAGGCTCGAGAGGGCAACAACATGTCCTCCAAACTTCC 558
QY 558 AAAGGGAGGCTTTTAAACAGATGTCAGAGCTCCCAACGAAGTCTGCTCTTAAACATT 617
DB 559 AAAAGGAAGCTTTTCGACAGATACGCTGAATGCCAAACGAAGTAGTCTGTACAGTACTT 618
QY 618 TTCAGGATTTGTAATACATCCGCGACAGAAATAGTGTAGAGATAGCGTGTCTCTGC 677
DB 619 TCCAAACGCTGTAGGCATTTCTCAAGAAATGTTACACGGGAAGAGTGTATGCTATTGCTTGC 678
QY 678 ACAGTTTGTATGATATATTCCTGTGTCATGAGTTTGAGCTGCGTTAAATCTAAGAAATATAC 737
DB 679 ATAGTATATACGATATACCTTGCAGCAGATGTCGCGCGGCACCTGCTGAGAAAGAAATGAC 738
QY 738 ATGTATGTTATGACGCTTCCATTTTGGCAGAAGCATTTATTAAGACAGACGAGGTTA 797
DB 739 ATGTATGTTATGCGCTTTCCACTTTTCCGAGAAATTTACTTCTCGAAGATTCACAGTCA 798
QY 798 CGCTTAATGAATAGGCGCACTTTTCAAAGAGAGGTGATGATTTCTTTTCTTTTCTTTG 857
DB 799 ACCTCGATGAGTCAATGTCATGTTTCCAAAGAGATGAGACAGGTGACCTTTTCTCTTTC 858
QY 858 CTGATGAAGTACTTTTAAATTAATAGTCAATAAATACAAATAATCTTGCATTTAGTTA 917
DB 859 CATCTGAGTACTCTTAAATTAATAGTCAATAGTTAATCTTAATATTTTAAAGTATTTGCA 918
QY 918 AATCTTACTTTCTCTCTCTAGTAGAATGTTTACTTTAAGGAATTTTAACTCACTAGG 977
DB 919 AACTTACTTTCCAGCTCTTAATAGAGGTTTACATGAAGGATTTTAACTAAGTATAG 978
QY 978 TTAATACTTGGTTTGTAAATTTACAAAGTAGATACCTATATTTCTGTACAAGAGTTA 1037
DB 979 TTAATACTTGGTTTGTAAATTTTCTAGAAATAGATATCTTCTTATTTGACAAAGGTAG 1038
QY 1038 GACAGTGTAGGTGTAGTGTATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1097
DB 1039 CGATTAAGGTGTAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1098
QY 1098 AGAAACCTTTGGCATGTTCAACACTGAAAGAGCAATCTTTTAGAGACACGGCTTCGGTTA 1157
DB 1099 AAAAGACTCTTGGATGTGCAACAGTGAAGAAATCTTGTAGAGATTTCTTCACTAGTTA 1158
QY 1158 ACTTTTGGTTTCCCTAAGATGAAGACATGATGATGATGATGATGATGATGATGATGATGATGAT 1217
DB 1159 ATTAAGTTTCCAAATGAGGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1218
QY 1218 CCAGCAAAAGATGACAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1277
DB 1219 AGACTAGTAAAGAACACGCAAGAGGCTTTAGTTTCAAAGGATTTGTTTATACAGTGT 1278
QY 1278 TTAATCATATCAGAAATATCAAGCAAGGCTTTAACTTACCAAGAGTATTTCTTTTCG 1337
DB 1279 TAAATCACTTGTACGTACAGGCAAGGCTTTACTTACTCAAGCTGTATTTCTTTTCG 1338
QY 1338 TGGAGTCTAATGATCCCGGTGATATCAATGGTGTACTGCTAGGTCTGGAATGGGATG 1397
DB 1339 TCGAATCAATTCGTTTCGAGAGTATCATTAACGGGTTACTGCCAGGTCTGAGTGGGATG 1398
QY 1398 TAGATAAGCAATTTCTCAACCTTTGTAATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1457
DB 1399 TCGATAATCATTTATTAACAGTCTTGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1458
QY 1458 CGCTTCAAGACGATATAGTAAATGGAAGTTTTCGTTCTGGTATGATGATGATGATGATGATGATGATGAT 1517
DB 1459 TTCTGAAAGACGATCTTTTGAATAGCAAGTTTGCACTTGGACCAAAACTGTCTCACAC 1518

QY 1518 TTATTTGGGATGAGGTGGGCAAAATTTTGGAAACGTTTCCCACTATCATCAAGAGAGAT 1577
DB 1519 ATGTGTGGGATGAGATTTCCCTAGCTTTCGGCAATGCTTTCCCATCGATCAAGGAAAGAT 1578
QY 1578 TGGTGAGCAGGAAATTTCTGGATGTAAGTGAGATGCTCTGAAGATCAAGATCCAGATCCAGATC 1637
DB 1579 TGATAAACCGAACTGATCAAAATTTACGGAGATGCGTTAGAGATCAGGTTGCCGATC 1638
QY 1638 TGTATGTACATGGAAGACAGGTTTCGTAGCTGAATACACAAAGCTGTGAGAGTTACCGC 1697
DB 1639 TTTATGTCACTTTCCATGATAGGTAGTTTCTGAGTACAAATGTCAGTGCACATGCGG 1698
QY 1698 ATCTAGATATCAAGAGGCTTAGAAGAGCTGAGCAAAATGTAAGGACGCTTATCAGAT 1757
DB 1699 TGTAGACATTTAGGAAAGATGGAAGAACTGAGGAAATGTAACAATGCACTGTCCGAAC 1758
QY 1758 TATCTATCTTAAGGCTGTGATAATTTTCGATATCGGAAGTTCAAGACATGTGCAAGG 1817
DB 1759 TGTCTGTACTTAAATTTCAACAGATGTCGATGTTGATGTTTTCCTCAGATGTGCCAT 1818
QY 1818 CTTTATAGTGTACTCTGTATGTCGACAGCAGATTAATCGTTGAGTGGCGGAGAAATAGAA 1877
DB 1819 CTTTAAAGGTTGATCCAAATGACTGCAGCAAGGTAATAGTAGCAGTTATGAGCAACGAGA 1878
QY 1878 GCGGTTTAACCTTACTTTTGTATGAAGCCACCGAGGAGATGTGGCTTAAGGCTTTAAA- 1936
DB 1879 GTGGTCTTACTCTCACGTTTGAACAGCCCAAGCTAATGTTGGCTAGCATTTCAAG 1938
QY 1937 ---- AGCACGGCTGTGAGGCGGTGTATGTTGAACCGACATCCGAAGAGGTGAACG 1991
DB 1939 ATTCTGAAAGGCTTCTGATGGGCGTTGGTAGTTACCTCAAGAGATGTTGAGGAACCGT 1998
QY 1992 TAAATAAATTTCTATTTGCTGAAAGAGGAGATTTGCTGTGTGTGCGAAGAAAGTCACTGTT 2051
DB 1999 CCATAAAGGTTGATGGCCCGGTGTAGTTACAAATTTGGCCGATATCTGGGACGCTC 2058
QY 2052 TGAGGATGCTTAATTTAGAGCAGGAGTTGAGTCCCTCAAGATTTCCATTAAGGCTT 2111
DB 2059 CTGAATCTTATACACTAGGAGCGGAGATTTGAGTCTCTCGAGCAGTTTCTATATGCA 2118
QY 2112 GCGTGTAGTGTGATTACAAAGCAATGSCATCGGTTGCTACACTGGCTCACTCAAG 2171
DB 2119 CAGTGTGTTAATTTATAGCAGATGTTTCGATTCGATGTCGTCCTCTGCTGCGGTGCGA 2178
QY 2172 TTCAACAAATGAAGAACTATGTGAGCAGTTTGGCAGCTTCGTTGTCCGCCACTGTATCAA 2231
DB 2179 TTCAACAAATGAAGAACTTTATAGACAGCTGTGTAGCTGCTCTGCTGCGGTGCGA 2238
QY 2232 ATCTATGCAAGTCACTAAAGGATGAAGTGGGTATGATTTCTGATTCAGGGAGAAAGTTG 2291
DB 2239 ATCTAGTGAAGATCTTAAAGATACAGCCGCGATTTGACCTTGAACCTCGTCAAAAGTTCG 2298
QY 2292 GTGTTTGGGATGTCATTTTGAAGAGTGGCTCTCAACCTCGCGCAAGGTCATTCAT 2351
DB 2299 GAGTTCTGAGTGTGCTTGAAGAGTGGCTAGTTAAACCTCCGCAAGAACCATGCAAT 2358
QY 2352 GGGGAGTGTCTGGATTACAAAGGAAATGTTTACTGCACTTTCTATCTTATGAAGGAG 2411
DB 2359 GGGGAGTGTGTAGACTCATGCGAGGAATATCACTGCGCATTTACTGGAGCAGATGAAT 2418
QY 2412 ATGAATGTGTAGAGCGACTGAGGAGGTTGGTGTATCATCTGTATCAATGTGTAT 2471
DB 2419 TTGGCATTTTACGTGCGATTAATGCGACGGTGGCTGTGAGTTTCTGAGTTCGCTGTAT 2478
QY 2472 ATTCTGATATTCAGAGCTCCAAATCTCGAGGAACAAATGAGAGCGGTGAACCCACG 2531
DB 2479 ATTCTGATATGGCTAAACTCAGGACTCTTGAGAAAGATTTGCTCAAGAGTGGAGAACCAACG 2538
QY 2532 AACCTACTGCAAGATGTTACTTGTGATGGGCTGCTGTTGTTGGAAGTGTACAAAGGAG 2591
DB 2539 TTAGTTTCAGCAAGAGTGGTTTGTGATGTCGCTTCCAGGGTGGGGAAGCAAGGAAA 2598
QY 2592 ATTTGAAAGATTTGATTTGATGAGGATTTGATTTGTTGTTCTCTGGAAACAAAGCTGCTG 2651

Db 2599 TTCTTTGAGAGTAAATTTTGAAGAAGATCTAATTTCTGTCTCTGTGCTGCTCAAGCTGCCG 2658
Qy 2652 CTATGATCAGAAGAGGGCTAAATTCATCTGAGTGTATGAAGAGCCAAATGCAATGTGA 2711
Db 2659 AGATGATCAGAAGAGAGCTAAATGCGTCGGGCATAATAGTGGCTACAAAGGATAATGTGC 2718
Qy 2712 GAACGGTAGAATTCACCTTCTAATGATCCAAAACCGCGATCAC-----ACAAGAGGC 2762
Db 2719 GCACCGTCGATTCATTTTGTGATGAATTAACGGGAAGGGCAACGCTGTCAGTTCAAAAGAT 2778
Qy 2763 TTTTATATGATGAAGGGTGTAGCTGCACACCGGTTGTGTAACTTCTCTGGTGTATCT 2822
Db 2779 TGTTCATAGACGAAGGTTTGTATGTCATATCTGTTGTGTGAATTTCTTGGTTGAAATGT 2838
Qy 2823 CTGGTTGCGACATCGCATACATTTACGGAGATACACAGCAGATTCCTTTCAATTAACAGAG 2882
Db 2839 CTCTGTGCGATATGCGATATGTTTATGGAGACACCCCAAAATTCGCTATCATCAACAGAG 2898
Qy 2883 TTCAAGAAATTTCCGGTATCCCAAAACATTTTGAAGAGCTGCAAGTGGATGAATTTGAGATGA 2942
Db 2899 TAACTGTGTTCCCGTACCTCTGCGCACCTTTGCAAAATTTGGAGGTGCGACGAAGTCGAACAA 2958
Qy 2943 GGAGGACCACTGAGATGCCAGGTGATGTGAATTTTTTCTTACAAATCGAAGTACGAAG 3002
Db 2959 GAAGAACTACTCTTCTGCTGCTCGGCTGATGTCAACACTTCTCTAAATCAAAGGTATGAAG 3018
Qy 3003 GAGGGTGACAAACACTTCACTGTACACGATCGGTCTCATCTGAGATGATAGCGCGTA 3062
Db 3019 GACACGTAATGTGCACGTCTTCTGAAAGAAATCAGTTTCCAGGAATGTTAGTGGGG 3078
Qy 3063 AGGAGTACTAAACAGATGTTTCCAAACCACTTAAAGGGAAATTTGTAACCTTCACTCAGG 3122
Db 3079 CTGGCTCTACTCAATCTGTGTCGAAGCGCTTAAAGGGAAATTTTGACITTCACACAGT 3138
Qy 3123 CTGATAAATTTGAGTTAGAGAGAGGGCTTATAGAATGTGAACACCGCTTCATGAGATCC 3182
Db 3139 CTGACAAAGGAGGCGCTTCTCTCAAGGGGCTACGCAGATGTCCATCTGTATCATGAGGTAC 3198
Qy 3183 AAGGAGAAACCTTTGAAGATGTGTCGTGTCAGATTCAGGCAACTCCACTGACTCTGA 3242
Db 3199 AAGGTGAGACTTATGACAGAGTATCGTTAGTTCGACTTAAACCTACCGCTGTATCTATCA 3258
Qy 3243 TTTTCAAAGTCTTCCCGCATGTTCTAGTTCGCTCTGACTAGACACACAAAGAGCTTCAAT 3302
Db 3259 TCGCAAGAGACAGTCCGATGTTCTGTCTCTGTTGTCAAGACACACAAATCCCTAAAGT 3318
Qy 3303 ATACACCGTAGTGTAGATCCCTTTAGTACAGATAATTAGTGAATTTGTTCTTTAAGCT 3362
Db 3319 ACTACACCGTTGTATGGATCCCTTTAGTTAGTATCATTTAGAGATTTTGAACGGGTTAGTA 3378
Qy 3363 CCTTCTCTTTAGAAAATGTATATGTAGAACGAGGTAGTAGTACAAATTCAGATGGATG 3422
Db 3379 GTTACTTATTAGACATGTAAAGTAGATGACAGGTACTCAATAGCAATTCAGGTGCACT 3438
Qy 3423 CAGTGTTCAAAGGTCAATAATCTTTGTGGCAACACCTTAATTCAGGAGACTTTCCAGATC 3482
Db 3439 CTGTGTTTAAAAATTTTCAATCTTTTGTAGCAGCTCCAAAGACTGGAGATATATCTGATA 3498
Qy 3483 TACAGTCTATATAGATGATGCTCCCTGTTATAGTACTATATTAACAAAGTATGATG 3542
Db 3499 TGCAATTTTACTATGATAAGTGTCTTCTGCGGAACAGCACGCTGTTTGAACAACTACGACG 3558
Qy 3543 CTGTTTACCATCAGGTTTACGTATATAGTCTTAAATGTGAAGGATTTGTTCTTCAATTTT 3602
Db 3559 CTGTTTACATGAAATTTGACTGACATTTCTCTGAATGTCAAGATGTGATATAGATATGT 3618
Qy 3603 CCAAAAGTATTCGATGCAAGAGAGGTGAACCATGTCTAGAGCCAGTTTTGGCTACCG 3662
Db 3619 CTAAGTCTGTAGTGTCTCCGAAGAGATGTCAACCAACTTTTAAATCCGATGTTACGAACG 3678
Qy 3663 CGGCGGAACCGGCTGAGGACTACTTCGAAATCTGTTGCAATGATTAAGAA 3722

Db 3679 CGGCAGAATGCTCCCCAGACTGGACTGTTGGAAAATCTAGTTGCGATGATTAAGAAGAA 3738
Qy 3723 ATTTCAACGCACACAGACTCTGACGGGACGATTTGACATTTGAGAGCACCGCATCTGTTGTAG 3782
Db 3739 ATTTTAATTCACCAGAGTTGTCGGAGTAGTTGATTTGAAATAATCTGCACTCTTTAGTGG 3798
Qy 3783 TAGATAAGTTTTTTGTATAGCTATTTTATTAAGAAAAGAAAATACACAAAATAATTTGCTG 3842
Db 3799 TAGATAAGTTTTTTGTATAGTTATTTTAAAGGAAAAGAAAACCAAAACAAAATTTTTT 3858
Qy 3843 GAGTATGACGAAGGATTCATGATCAGATGTTTGGAAAACAGGAAGAAGTACTATTGG 3902
Db 3859 CACTGTTTAGTAGAGAGTCTCTCAATAGGTGGATAGCAAGCAAGCAAGTCACAATTG 3918
Qy 3903 ACCACTTGGCTAACTACATTTTACAGATCTGCGGCCATTCGATCAGTACAAAGCACAAGA 3962
Db 3919 GTCAGTTGGCCGATTTTGTGATCTTCCAGCCGTTGATCAGTACAGGCATATGA 3978
Qy 3963 TCAAGGCTCAACCAAAACAGAAATTTGACCTTTCAATTCAGATGATACCTGCTCTGC 4022
Db 3979 TTAAGCGCAACCGAAGCAGAAA CTGGATCTGTCNAATTCAGACAGAAATATCCAGCGTTC 4038
Qy 4023 AACAATTTGCTACCATTCGAACAGATCAACGGTATTTTGG-- -CGGTTTCTCAGAGC 4079
Db 4039 AAACGATTTGTGATCATTTCAAGAAAATCAACCAATTTTGGTCTCTTTTCAGTGAGC 4098
Qy 4080 TTACAAGTTGCTGCTCGAGGCAATTTGATTTCTAAGAAGTTTCTTTTCTTTTACTAGGAAA 4139
Db 4099 TTACAAGGCAATTTACTTGACAGTATTGACTCAAGCAGATTTCTTGTCTTACGAGAAGA 4158
Qy 4140 CTCAGACAAGATTCAGAAATTTTCTCGGATCTCGACTCGACGTTCCTATGGATGTT 4199
Db 4159 CACCGGCTCAGATCGAAGATTTCTTCGGAGATCTAGACAGTCTGATGTCCTCAATGACGTAC 4218
Qy 4200 TAGAATCTGGATTTCTTAAGTATGATAAGTCAAGAACGAGTTTCATTGTCTGTAGAGT 4259
Db 4219 TTGAGTTGGATGTTTCAAGTATGATAAGTCTCAAAACGAGTTTCATTGTGCTGTTGAGT 4278
Qy 4260 ATGAAATATGAAAAGATTTGGGCTCTCAATCAGTTTTCGCGCAAGTGTGGAACAAGGGC 4319
Db 4279 ACGAATCTCGAGGAGACTGGGCTCGAGGATTTCTTGGCAGAAAGTGTGGAACAAGGGC 4338
Qy 4320 ACAGAAAACAACTTTGAAGGATTCATTCGTAAGTCAAGACAGATGCTGTGGTATCAAA 4379
Db 4339 ATGAAAACCAACCTCGGAAAGATTTACACTGCTGATATAAAACGTTTATTTGGTACCGA 4398
Qy 4380 GGAAGCGGTGATGTGACTACTTTTCATCGGCAATCTGTTTAAATAGCAGCTTGTCTGG 4439
Db 4399 GAAAGAGTGTGATGTTTACAACTTTTATCGTAAATCCGTCATCATTTGCTTCGTGCTTG 4458
Qy 4440 GTTCAATGTTTACCGATGGAAAAGGTCTATAAAGGTGCTTTTGTGGAGAGCATTCGGTTT 4499
Db 4459 CATCAATGCTCCCGATGGAAAATTTGATAAAGAGGCTTCTGCGGAGATGACAGTTTGT 4518
Qy 4500 TGTATTTTCCAAGGGTTTGGATTTCCCTGACATTCAGTCAATGCTTAATCTCATGTGA 4559
Db 4519 TGTACTTTCTTAAGGGTTGTGAGTATCCCGATATACAAACAGCTGCGCAATCTAATGTGA 4578
Qy 4560 ATTTTGAAGGCAAACTCTATAGAAAAGAGTACGGTTACTTTTGTGGTAGATACATCATAC 4619
Db 4579 ATTTTGAAGCAAACTGTTCAAGAACCAATATGGGTACTTCTGCGGGAGGTACGTGATTC 4638
Qy 4620 ACCATGATAAGGGAGCAATAGTGTATTTATGATCTTTTGAAGTTGATCTCCAAACTTGGGG 4679
Db 4639 ATCAGATAGAGGTTGATAGTATCTACGACCTTTTGAAGCTGATTTTGAAGCTTGGTG 4698
Qy 4680 CAAAACATATCAGGATATGATCACTTAGAAGGTTTAAAGGCTCTTTGTCGATGTTG 4739
Db 4699 CTAACACATCAAGGATTTGGATCATTTTGGAGGAGTTTCAAGAGATCCCTCTGTGATGTTG 4758
Qy 4740 CTGTTCTCGTCGGAATACTGGTGTCTTAGGCTTTCCGCAAGCTGAAACGCACTATCAAGGAG 4799
Db 4759 CTGAGTCTGTT---GAAACAATTCGCGTATTACACAAATTTGACGACGCTGTGCGGAGG 4815

```
QY 4800 TTCAATAAACCGGATGTAGTGGTTCGTTTCGTTTAAATTTGTTGTAACAAATTTTTCGTG 4859
D 4816 TTCAATAAACCGGCCCCACCGTGGTTCGTTTAAAGAGTTTGTAGTTAAGTATTTGTCAG 4875
QY 4860 ATAAATTTTATTTAGAACTTTGTTTAAATGGCTGTAGTCTCAGAGATACGTCAAA 4919
D 4876 ATAAAGTTTGTAGAGTTTATTTCTGATGGCTCTAGTGTGTAAGGTAGGTAAAT 4935
QY 4920 ATTACGAGTTCATTGATCTTTTCAAAACAGGATGAGATACCTCCGGCATTCATGACTAAG 4979
D 4936 ATTAATGAGTTTATCGATCTGTCAAAAGTCTGAGAACTTCTCCGTCGATGTTCAAGCCT 4995
QY 4980 GTCAAGAGTGTAGATATACGATCTGACAGAGTATGCTCTTAAGATGATAGTCTT 5039
D 4996 GTAAAGAGTGTATGTTTCAAAAGTGTGATAAGATTATGCTCATGAAATGAATCATG 5055
QY 5040 TCTGATGTAGATTACTTAAAGGTGTTAAGTTAGTTTAAAGAAAGGTATGTGCTTAGCT 5099
D 5056 TCTGAAGTAAATCTCTTAAAGGTGTAAACTTTATAGAGGTGGTATGTTGCTTAGTC 5115
QY 5100 GATTGGTAGTGTCTGGGAGTGGAAATCTCCCGATAACTGCGGTGGTGTGTCAGTGT 5159
D 5116 GGTCTTGTGTGTCGCGTGAAGTGAATTTACAGATAATTTGCGGTGGTGTGAGTGT 5175
QY 5160 TGTATTGTAGATAAGAGATGAAGAGTGAAGCAACGCTGGGTGCTATCAGCC 5219
D 5176 TGCATGGTTGACAGAGATGGAAGAGCGGACGACCTGGGGTCTATATTACACT 5235
QY 5220 CTTGCTTGCAAAAAGAAATTTCTTTTAAGCTAATCCTTAATTTCAATTAACATCCGAG 5279
D 5236 GCTGCTGCTAAAGAGCGTTTCAAGTTAAGTGTCCCAATTTACGATTTACACAAAG 5295
QY 5280 GATGCTGAGAAAGCAGCGTGGCAAGTGTAGTGAATATCAAGAGAGTGGCTATGGAAGAA 5339
D 5296 GATGCAAGAAAGAACATATGCGAGTCTTAGTAAATATTAATAATGTAATAATGAGTGC 5355
QY 5340 GGATACGTCTTTATCTTTGGAGTGTGTTTCAATTTGTGTAGTACATAAAATATATGTA 5399
D 5356 GGCTACTGCTTGTGTCATTTAGAAATTTGTGCTGTGTGTTTATTAATAATAATATA 5415
QY 5400 AGAAAGTTTGTAGGGAACGATTTTGTAGTGTGACAGCGCTCGCAATTTGAACCTCACT 5459
D 5416 AATTTGGTTTGAAGGAGAAAGTAAACGAGTGTGACAGTGAGAGACCCATGGAACCTTCA 5475
QY 5460 GAAAGGTTTGTGAGGAGTTTCGTGGAATGAGTACCAATGCTGTGAACTCGAAAAGGTT 5519
D 5476 GAAGAAGTTGTTGATGAGTTTCATGGAGAAATGTTCCAAATGTCGTTAGACTCGCAAAGTTT 5535
QY 5520 CCGGAAACAAAGAAATGGTAGGTAAATGTTTAATTTAATTTAATTTAGGTAAAGGCGTTCA 5561
D 5536 CGAAACCAATTCCTCAAAAGAGGTCGGAATAATTAATAATTTAGGTAAAGGCGTTCA 5595
QY 5562 ---AAGAAATAATAACAGTGTGAAGAGGTTTAAATTTAGGAAATTTGAGGATAAT 5618
D 5596 GCGGAGAGGTCTAAACAAAGTTTGTGATGAAGTTGAAAGAGTTTGAATAATTTGATT 5655
QY 5619 GTAAGTATGACGAGTCTATCGCGTTCATCGAGTACGTTTTTAATCAATATGCTTATACAA 5678
D 5656 GAAGATGAAGCCGAGAGCTCGGTCGCGGATTTCTGATTCGTTAATAATATGCTTACTCAA 5715
QY 5679 TCAACTCTCGAGCCAAATTTGTTTATCTTCTTCGCTTACGAGATCTCTGTGAGCTGA 5738
D 5716 TCACCTCTCCATCGCAATTTGTGTTTGTTCATCTGTATGGGCTGACCCCTATAGAATTGT 5775
QY 5739 TCAATCTGTGTACAAATTCGATTCGGTAAACAGTTTCAACGCAACAGCTAGGCAACAG 5798
D 5776 TAAACGTTTGTACAAATTCGTTAGGTAAACAGTTTCAACACAGCAAGCAAGAACTACTG 5835
QY 5799 TCAACAGCAATTTGCGGATGCTCGAAACCTGTGCTAGTATGACAGAGTGAAGTTTCCGTG 5858
D 5836 TTCAACAGCAATTCAGCGAGGTGTGGAAACCTTTCCTCTAGAGCAGCGTCAGATTTCCTG 5895
```

```
QY 5859 CATCGATTTCTATGTGTATAGATATAATTCGACGCTTGTATCCGTTGATCAGCGCGTTAT 5918
D 5896 GCGATGTTTATTAAGGTGTACAGTACATGCACTTTTAGATCTCTTAATTTACTGCGGTGC 5955
QY 5919 TAAATAGCTTTTGATCTACTAGAAATAGAAATAGAGGTTGATATCAACCCGACCCGAATA 5978
D 5956 TGGGGGCTTTTCGATCTACTAGGAATAGAAATATCGAAGTAGAANAACAGCAGAGTCCGACAA 6015
QY 5979 CTACTGAAATCGTTAAAGGCTACAGAGGTAGACCATGCTACTGTAGCTATTAAGGCTT 6038
D 6016 CAGCTGAAACGTTTAGATGCTACCGCAGGGTAGACGACGCTACGGTTGCAATTCGGTCTG 6075
QY 6039 CAATCAATAATTTGGCTTAATGAACCTGTTTCGTGGAACCTGCAATGTTCAATCAAGCAGGCT 6098
D 6076 CTATAAATAATTTAGTTAATGAACCTAGTAGAGGTACTGCACTGTACAAATCAGAATACTT 6135
QY 6099 TTGAGACTGCTAGTGGACTTGTCTGACCACTCCCGCTACTTACTAGCTATTTGTGTGAG 6158
D 6136 TTGAAGATGATGTCGGGTTGGTCTGACCTCTGCACCTGTCATCTTAAATGTCATAGGTGCT 6195
QY 6159 ATTTCCTAAATAAGTCTGCTGAGACTTTAAATTTCAAGGTGCTGATACCAAAATCAGC 6218
D 6196 GAATATAAATTTTGTGTTTCTTAAACACACGCTGCTGATACGATAACGTAC 6247
QY 6219 AGTGGTGTGTTTCGTCCTTAAATATAACGATTTGTCATATCTGATCCAAACAGTTAAACCA 6278
D 6248 AGTGTGTTTTCCTTCCCTTAAATCGAAGGTAGTGTCTTGAGCGCGCGGAGTAAACATA 6307
QY 6279 TGTGATGTTGATGCTGCTGATGCGTAAACATCGGAGAGGTTTCAATCTCCCTAA 6338
D 6308 TATGGT-TCATATATGTCGTAGGCACGTAAAAAGCGAGGATTCGAATTTCCCGGAA 6366
QY 6339 CCGCCGCTAGCGGCCCA 6355
D 6367 CCGCCGCTAGCGGCCCA 6383

RESULT 12
TMVK2XX
LOCUS
DEFINITION
Tobacco mosaic virus (K2 strain), complete genome.
ACCESSION
292909
VERSION
292909.1
KEYWORDS
coat protein; complete genome; RNA-dependent RNA-polymerase.
SOURCE
Tobacco mosaic virus
ORGANISM
Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE
1 (bases 1 to 6383)
Belenovich, K. V., Genozov, E. V., Novikov, V. K. and Zavrilev, S. K.
Properties and genome structure of Tobacco mosaic virus K2-strain
Unpublished
REFERENCE
2 (bases 1 to 6383)
Genozov, E. V.
Direct Submission
Submitted (13-MAR-1997) Genozov E. V., Institute of Agricultural
Biotechnology, Laboratory of Molecular Virology, 42 Tsimiryaevskaya
st., Moscow, Russia, 127550
Location/Qualifiers
1. .6383
/organism="Tobacco mosaic virus"
/mol_type="genomic RNA"
/strain="K2 (Kazakhstan strain)"
/db_xref="taxon:12242"
72. .4922
/function="RNA-dependent RNA polymerase"
/note="ORF2, readthrough protein"
/codon_start=1
/product="183K protein"
/protein_id="CA07439.1"
/db_xref="GI:1890235"
/db_xref="GOA:P89676"
/db_xref="Swiss-Prot:P89676"
source
CDS
```

```

/translation="MATQTATSSALLETVRGNNTLVNDLAKRLDYDTAVDFNARDR
RPKNVFSVSEETLIATKAYPEFOITFYNTQNAVHSLAGGLRSLELYLMMQIPYG
SLTYDIGNFASHLFGKRAYVHCMPNLDVDRIMRHEGQKDSIELYLSRLERGNKHVP
NFKQAEFAERFAEMNEVCHDTFOTCRHSQECYTVGVYALHSHIYDIPADEFGAALL
RKNVHCYAAAFHSENVLLDSDSHVNDIENACFORDCGRDLTTFPASESTLNHSHYSN
ILKVYCKTYPASNREYVMEFLVTRVNTWFCFSRDTFLFKGVAKHGVDSQEFYK
AMEDAHYKKTLCMNCSEIRLLEDSSVNTWFPKORDMVIIPUDISLTSKRTKEV
LVSKDFVYTLNLRHTYQAKALTVSNLVSFVESIRSRVVIINGVTARSEMVDKSLLOS
TASGLIHLQWCSIVYTGPKVQOMKNFIDSLVASLSAAVSNLVKILKDTAIDLETRO
KFGVLDVASRKLWLPKSAKHANGVETHARKYHVALLEHDEFIITCDNWRVAVSS
EVSUYDMAKRLTRLLKDGEGEHPVSKAVLVDGPGCGTKTEILSRVNFEDLILV
PGROAAEMRRRANASGIIIVATKONVRTVDSFLMNYCKGAKCOFKRFLIDEGMLHTG
CVNLFVMSLCDIAVYGDVTOQIPYINRVTFPYPAPFAKLEVDVETRETTLCPCAD
VTHFLNQRVGHVWCTSSSEKSKSQEMVSGAASINPVSKPLKGKILFTQSDKEALLS
RGVADVHTHEVQGETYADVSLVRLTTPVSIIRADSPhVLVSLSRHTKSLKYTVVM
DPLVSIIRDLERVSYLLDMYKVDAGTQXQLOVDVFPKFNFLVFAAPKTDGIDSMOPY
YDKCLPGNSLLANNYDAVTWKLTDIISLVNKCOCILDMSKSVAAPKDVKPTLIPMVRTAA
EMPROTGLLENLAVMIKRNFNSPELSGVDIENTASLVVDKFFDSYLLKEKRPKNF
SLPSESINRWIAKQEOVTIGOLADFPVDLPAVDYRHMIAKQPKQLDLSLOTEVP
ALQTIYVHSHKINIAFGLSELTROLLSDISDRFLFTRKTPAQIETDFGDLDSHV
FMVLEVDYKDKSQNEFHCAVEIYIWRRLGLEDFLAEVWKQGRKXTLLKDYTAGIK
TCLWYQRKSGDVTFFIGNVTIIASCLASMLPMEKLIKAFCDGDDSLLYFPKGEYDPI
QOANLNMNFEAKLFFKQYGVFCGRYVVIHHRGCIYVYDPLKILSLGARHAKIDWHL
EFPRLSCDVAESLNNAYYQLDDAVGEVHTKTPGGSFVYKSLVLYSKVLFRSLF
LDGSSC"

```

CDS

```

72..3422
/function="protein containing methyltransferase and
RNA-helicase domains"

```

```

/codon_start=1
/product="126K protein"
/protein_id="CAB07438.1"
/db_xref="GI:1890234"
/db_xref="GOA:P89676"
/db_xref="Swiss-Prot:P89676"
/translation="MATQTATSSALLETVRGNNTLVNDLAKRLDYDTAVDFNARDR
RPKNVFSVSEETLIATKAYPEFOITFYNTQNAVHSLAGGLRSLELYLMMQIPYG
SLTYDIGNFASHLFGKRAYVHCMPNLDVDRIMRHEGQKDSIELYLSRLERGNKHVP
NFKQAEFAERFAEMNEVCHDTFOTCRHSQECYTVGVYALHSHIYDIPADEFGAALL
RKNVHCYAAAFHSENVLLDSDSHVNDIENACFORDCGRDLTTFPASESTLNHSHYSN
ILKVYCKTYPASNREYVMEFLVTRVNTWFCFSRDTFLFKGVAKHGVDSQEFYK
AMEDAHYKKTLCMNCSEIRLLEDSSVNTWFPKORDMVIIPUDISLTSKRTKEV
LVSKDFVYTLNLRHTYQAKALTVSNLVSFVESIRSRVVIINGVTARSEMVDKSLLOS
LSPWFLHTVHLTKDILLISKALGPKTVSQHWDBEISLAGNAPSPISKERLNRKL
IKITENALEIRVPDLYVTFHDLRVSEYKMSVDMPLDIRKMBEITEMYNALSLSV
KNSDKPDVDFVSCQSLVDPMTPAAKIVIVAMNSNGLTLTFEQPTNEVALALQDS
EKASDGLAVTSDRVEPSIRGSMARGELQAGLSDGVPBSSYTRSEELSLQFHMA
TAGSLIHLQWCSIVYTGPKVQOMKNFIDSLVASLSAAVSNLVKILKDTAIDLETRO
KFGVLDVASRKLWLPKSAKHANGVETHARKYHVALLEHDEFIITCDNWRVAVSS
EVSUYDMAKRLTRLLKDGEGEHPVSKAVLVDGPGCGTKTEILSRVNFEDLILV
PGROAAEMRRRANASGIIIVATKONVRTVDSFLMNYCKGAKCOFKRFLIDEGMLHTG
CVNLFVMSLCDIAVYGDVTOQIPYINRVTFPYPAPFAKLEVDVETRETTLCPCAD
VTHFLNQRVGHVWCTSSSEKSKSQEMVSGAASINPVSKPLKGKILFTQSDKEALLS
RGVADVHTHEVQGETYADVSLVRLTTPVSIIRADSPhVLVSLSRHTKSLKYTVVM
DPLVSIIRDLERVSYLLDMYKVDAGTQ"
4906..5700
/function="movement protein"
/codon_start=1
/product="29K protein"
/protein_id="CAB07440.1"
/db_xref="GI:1890236"
/db_xref="GOA:P03584"
/db_xref="Swiss-Prot:P03584"
/translation="MALVIVKGVKNINEFIDLSKSEKLLPSNFTPVKSVWVKDKIMV
HENESLSEVNLKGLVIEGGYVCLVGLVSVGEWNLPCNCRGGVSVWVKDKRBERADE
ATLGSYYTAAAKRFPQVPVNTYITTKDAEKNIQVWLVNKNVKSAGYCPILSLFV
SVCTIVYKNNTKILGRREKTVSVNDGGPMELSEEVDEPMENVPMSVRLKPRKTSKRG
PKNNNLKGRSGRPKPKSFDEVEKEFDNLIDEAETSIVADSDSY"
5703..6182

```

CDS

```

/function="coat protein"
/notes="ORF4"
/codon_start=1
/product="17.7K protein"
/protein_id="CAB07441.1"
/db_xref="GI:1890237"
/db_xref="GOA:P89677"
/db_xref="Swiss-Prot:P89677"
/translation="MSYPTSPQKVFVLSLWVADPIELLNVCTNSLGNQFQTOQARTT
VOKQSEVWKPQSPQSTVPFPGDVYVYVNAVLDPLITALLGTGDFTRNSIIEVENROS
PTTAETLDTARRVDDATVAIRSAINNLVNLVRLGTGLYQNQTFESMSGLVWTSAPAS"

ORIGIN

Query Match      41.2%; Score 2617.4; DB 14; Length 6383;
Best Local Similarity 64.2%; Pred. No. 0;
Matches 4091; Conservative 0; Mismatches 2231; Indels 55; Gaps 8;

Qy 18 CGACAACAACATTAACAACAACAACAACATATTACAAACAACAACAACAACAATGGCAC 77
Db 23 CAACAACAACAACAACAACAACATTTACATTC-----TACAACATCAACATGGCAT 78
Qy 78 ACATACAATCTATAATTAGCAACGCCCTCTCTCAAGCGGTGAGTGGTAAAAACAACCTCG 137
Db 79 ACACAACAACAGCCACATCGTCGCTTGTGAGCCGTCGAGGTAAACATACCTTGG 138
Qy 138 TTAATGACCTTGCAGAAGCGCATGTACGATACGGCCGCTGGGAAGAAATTTAAACCCCGCG 197
Db 139 TCAACGATCTTGCAAAAGCGCGCTCTATATGACACAGCGGTGCGATGAATTTAATGCTAGG 198
Qy 198 ACCGTAGACCAAGAGTCAACTTTTCCAAACTATTAGCGNAGAGCAACACGCTTCTAGTCT 257
Db 199 ACCGCGAGGCTTAAAGTCAATTTTTCAAAAGTAGTAGCGGAAGAACAGACGCTTATTGCAA 258
Qy 258 CCAACGCGTACCCGAGGTCCAGATTACCTTTTATAATACTCAAAATGCCGTACACAGTT 317
Db 259 CCAAAGCTTACCAGAAATTCAAATTTACATTCTACACACGACGACGCTGTGCATTTCCC 318
Qy 318 TGCTGGAGGTTTGAGAGCATTAGAATTTGGAATATCTGATGTCTACAAGTTCCCTATGGAT 377
Db 319 TTGAGGCGGTCTCCGATCATTAGAATTGGAATTTGGAATATCTGATGTGCAAAATTTCCCTA 378
Qy 378 CGCGGACATATGATATAGTGGGAACTTTGCAGACATATTTGTTTCAAAGCGGAGGATACG 437
Db 379 CATTGACATATGATATCGGAGGTAAATTTTGCATCTCATCTGTTTCAAAGGCGGAGCATAC 438
Qy 438 TGCATTGCTGTATGCCCAATCTCGACATACAGATATAATAGCGGACGACGAAGCAAAAGG 497
Db 439 TTCACCTGCTGTATGCCCAATCTGATGTCCGCGACATATATCGGCGACGAGGCGCAAAAGG 498
Qy 498 ACTCAATTGAGATGTATTTGTCCAGATTGTCTCGTTCTTAAAGGTAATTTCTTGAGTTTC 557
Db 499 ACAGTATTGAACATATACCTTTCTAGGCTCGAGAGGGGCAACAAACATGTCCCAAACTTCC 558
Qy 558 AAAGGAGGCTTTTAAACAGATATGACAGAGCTCCCAACGAGCTCTGCTGCTCTTAAACTT 617
Db 559 AAAAGGAAGCTTTTCAAAGATACGCTGAAATGCAAAACGAAGTAGTCTGTGCAAGATAC 618
Qy 618 TTCAGGATTGTCCGAATACATCCGCCAGAGAAATAGTGTAGAGATACGCTGTGTGCTCTGC 677
Db 619 TCCAAACGCTAGGCATTCTCAAGATGTTACACGGGAGAGGTGTATGCTATTGCTTTC 678
Qy 678 ACAGTTTGTATGATATATTCCTGTGCATGAGTTTGGAGCTGGGTAAATCTTAAAGATATAC 737
Db 679 ATAGTATATACGATATACCTTCCGCGAGAGTTTCGCGCGGCACCTGCTGAGAAAGATGTAC 738
Qy 738 ATCTATGTTATGACGCTTCCATTTTGGCAGAACATTTATTACTAGACCACGAGGTTA 797
Db 739 ATGTATGTTATGCGCGTTTCCACTTTTCCGAGAAATTTACTTCTCGAAGATTTACACGTCA 798
Qy 798 CGCTTAATGAAATAGGCGCAACTTTTCAAAGAGAGAGTGATGATGTTTCTTTTCTTTT 857
Db 799 ACCTCGATGAGATCAATGATGATGTTTCCAAAGAGATGAGACAGGTTGACTTTTTTCTTT 858

```

Qy	858	CTGATGAAGATGACTTTTAAATATATAGTGCATAAATACAAAATATCTTGCATATGTAGTTA	917
Db	859	CATCTGAGAGTACTCTTAAATCATAGTCAATAGTTATCTTAATATCTTAAAGTATGTTTGC	918
Qy	918	AATCTTACTTTTCGCTCTAGTAGAATAGTTTACTTTTAAAGGAATTTTTAGTCACTAGGG	977
Db	919	AAACTTACTTCCCGAGCTCTAAATAGAGAGGTTTACATGAAGGAGTTTTTATGTAACTAGAG	978
Qy	978	TTAATCTCTGGTTTTGTAAAAATTTACCAAAGTAGATACCTATATTTCTGTCAAGAGTGTTA	1037
Db	979	TAAATACCTGGTTTTGTAAAAATTTCTAGAAATAGATACTTTCTTATTTCTACAAGGTGTAG	1038
Qy	1038	GACAAGTAGGGTGTGATGATGATCAGTTCTATAGAGCGATGGAAGACGCTTTGCTTTTACA	1097
Db	1039	CGCATAGGGTGTAGATAGTGAGCAGTTTTACRAAGGCTATGGAAGACGATGGCACTACA	1098
Qy	1098	AGAAAACCTTGGCCATGTTCAACACTGAAGAGCAATCTTTAGNAGACACGGCTTCGGTTA	1157
Db	1099	AAAAGACTTTCGGATGTGAACACAGTGAAGAATCTTTGTAGAGGATCTTTCATCACTAGTTA	1158
Qy	1158	ACTTTTGGTTCCCTAAGATGAAGACATGGTGAAGTACCGCTGTTTGAAGGTTCTTATTA	1217
Db	1159	ATTACTGGTTTCCAAAATGAGGATATGTTAGTATTTCCACTATTTTGAATATCTCTCG	1218
Qy	1218	CCAGCAAAAAGATGACGAAGGATGAGTCAATGTTTAATCGTGACTTCGTTTACACAGTGC	1277
Db	1219	AGACTAGTAAAAAGAACACGAAGAGGCTTTAGTTTCAAAGGACTTTGTTTATACAGTGT	1278
Qy	1278	TTAATCATATCAGAACATATCAAGCCAAAGGTTTAACTTTACCAGAAAGTATATCTTTTCG	1337
Db	1279	TAAATCACTTCGTAACGTACAGGCCAAAGCGCTTACTTACTCCAAAGCTGTTATCTTTTCG	1338
Qy	1338	TGGAGTCTATAAGATCCGGGTGATAATCAATGGTGTGTACTGCTAGTGTCTGAATGGGATG	1397
Db	1339	TCGNATCAATTCGTTTCGAGAGTGATCAATTAACGGGGTACTGCTAGTGTCTGAGTGGGATG	1398
Qy	1398	TAGATAAAGCAATTCCTTCAACCCCTTGTCAATGACTTTCTTCTTTCAGACATAAGCTGCGTG	1457
Db	1399	TCGATAAATCAATTAATACAGTCTCTGTGATGAGCTTCTTCTTCTACATACCAAGCTTCGCG	1458
Qy	1458	CGCTTCAAGACGATATAGTAATGGGAAAGTTTCGGTCTGGATAGAACACCACTTCTGAAC	1517
Db	1459	TTCTGAAGAAGCGATCTTTTGTGATTAGCAAGTTTGCACTTGGTCCAAAACATGTCTCAACAC	1518
Qy	1518	TTATTTGGGATGAGTGGGCAAAATTTTTTGGAAACGTTTTTCCCCACCTATCAAAAGAGAGAT	1577
Db	1519	ATGTGTGGGATGAGATTTCCCTAGCTTTTCGGCAATGCTTTTCCCATCGATCAAGGAAGAT	1578
Qy	1578	TGTTGAGCAGGAAAATTTCTGATGTAAGTGAGATGCTCTGAAGATCAAGATCCAGATC	1637
Db	1579	TGATAAACCGGAACCTGATCAAAATTTACGGAGATGCGTTAGAGATCAGGGTGCCTGATC	1638
Qy	1638	TGTATGTCAATGGAAAGACAGGTTTCGTAGTGAATFACCAAGTCTGAGGAGTTACCGC	1697
Db	1639	TTTATGTCACTTCCATGATAGGTTAGTTTCTGAGTACAAAATGTCAAGTGGACATGCGCG	1698
Qy	1698	ATCTAGATATCAAGAAGGACTTTAGAAGAAGCTGAGCAAAATGTACGACGGTTATCAGAAT	1757
Db	1699	TGCTAGACATTAGGAAAAGANTGGAGAAACTGAGGAAATGTACATGCAATGCTGTCGAC	1758
Qy	1758	TATCTATCTTTAAGGGTCTGATAATTTTCGATATTCGCGAAAGTTCAAAAGACATGTGCAAGG	1817
Db	1759	TGTCGTGACTTAAAAACTCAGACAAAGTTTCGATGTTGACGTTTTTTTCCAGATGTGCCAAT	1818
Qy	1818	CTTTAGATGTTAGTCTCGATGTGTCAGACAAGAGTAATCGTTTGCAGTGGCGGAGATAGAA	1877
Db	1819	CTTTTGAAGTCCGATCCAATGACTCCAGCAAAAGGTAATAGTAGCAAGTTATGAGCAACGAGA	1878
Qy	1878	GCGGTTAACTCTTACTTTTGTATGATAGCCAAACCGAGGAGAAATGTGGCTTAAAGGCTCTTAAA	1936
Db	1879	GCGGTCTTACTCTCACGTTTGAACAGCCCAACCGAAGCTTAATGTTTGGCGTAGCATTTGCAAG	1938
Qy	1937	-----AGCACGGCGTCTGAGCGCGTGTATGCTCTTGAACCGCACATCCGAAGAGGTGAAAG	1991

1939 ATTCTGAAGAGGCTTCTGACGGGGCGTGGTAGTTTACCTCAAGAGATGTTTCAGGAACCGT 1998
1992 TAAATAAAATTTCTATTGCTGAGAAAGGGAGATTGCTGCTGTGTGCAGAAAAGTCATGTT 2051
1999 CCAATAAGGGGTTTCAATGGCCCGTGGTAGTTACAAATTTGGCCGGATTATCTGGCAGCTTC 2058
2052 TGACGAATGCTAACTTAGAGCAACGAGAGTTGGAGTCCCTCAACGATTTTCCATTAAGGCTT 2111
2059 CTGAATCTTTCATACACTAGGAGCGAGGAGATTGAGTCTCTCGAGCAGTTTTCATATGGCAA 2118
2112 GGTGGTAGTGTGATTTACAAAGCAATGGCATCGGTGTCTACACTGGCTCAGTCAAG 2171
2119 CAGCTAGTTCGTAAATTCATAAGCAGATGTGTTTCGATCGGTGTACACGGGCCCTCTTAAG 2178
2172 TTCAACAAATGAAGAACTATGTGACAGATTTGGCAGCTTCGTGTGTCGCCCACTGTATCAA 2231
2179 TTCAACAATGAAAACCTTTATAGACAGCCTGGTAGCCTCGCTCTCTGCTGGGTGTGCA 2238
2232 ATCTATGCAAGTCACTAAAGATGAAGTCGGGTATGATTTCTGATTTCCAGGGAGAAAGTTG 2291
2239 ATCTAGTGAAGATCCTTAAAGATACAGCTCGGATGACCTTGAACCTCGTCAAAAGTTTCG 2298
2292 GTGTTTGGGATGTCATTTTGAAAAAGTGGCTCTTCAAACTCGCGGCCAAAAGGTCAATTCAT 2351
2299 GAGTTCTGGATGTGCTTTCGAAAAGGTGGCTAGTTTAAACCATTCGCAAGAAACCAATCGAT 2358
2352 GGGGAGTTGTCTCGATTTACAAAGGGAAAAATGTTTACTGCACTTCTATCTTATGAAGGAG 2411
2359 GGGGGTTGTTGAGACTATCGCAGGAAATATCACTGTCGCATTAATCTGGAGCAGCATGANT 2418
2412 ATAGAATGGTGACTGAGAGCGACTGGAGAGGGTGGCTGTATCATCTGATACAATGGTAT 2471
2419 TTGGCATTATTACGTGCGATAAATCTGGCAGCGGTGGCTGTGAGTTCTGAGTCGGTAGTAT 2478
2472 ATTCTGATTTGCAAAAGCTCATAAATCTGAGGAAAACAATGAGAGACGGTGAAACCCACG 2531
2479 ATTCTGATATGGCTAAACTCAGGACTCTCGAAGAGATTGCTCAAAGATGGAGAACCAACG 2538
2532 AACCTACTGCAAGAGATGTAATCTTGATGGGGTGCTGGTTGTGAAAGTACAAAGGAG 2591
2539 TTAGTTTCAGCAAGAGTGTTTTGTGTGATGGCGTTCCAGGTGCGGAAAACAAGGAAA 2598
2592 ATTTTGAAGATTTGATCTTTGATGAGGATTTGATCTTTGTTCTCTGGAAAAACAAGCTGCTG 2651
2599 TTTCTTTCGAGAGTTAAATTCGAGAAGATCTAAATCTTGTCTCCCGGTCTCAAGCTCGC 2658
2652 CTATGATCAGAGAGAGGCTAAATTCATCTGGAATGATTAAGAGCCACAATGAGACAAATGGA 2711
2659 AGATGATCAGAAGAGAGCTAAATCGCTCGGGCATTAATAGTGGCTCAAAAGGATAATGTGC 2718
2712 GACCGTAGATTCACCTTCTAATGCATCCAAACCGGATTCAC-----ACAAGAGGC 2762
2719 GCACCGTCGATTCATTTTGTATGAATTAACGGGAAGGGGCACTGTCTGAGTTCAAAGAT 2778
2763 TTTTATTTGATGAAGGGTTGATGTCGACACCGGTTGTGTAACTTAACTCTCGTGTCTATCT 2822
2779 TGTTTCATAGACGAGGTTTGTGCTGCATCTGTTGTGTGAAATTTCTTGTGTTGAAATGT 2838
2823 CTGGTTGGGACATCGCATACATTTACGAGATACACAGCAGATTCCTTTCAATTAACAGAG 2892
2839 CTCTGGCGATTTGCATATGTTTATGGAGACACCCAAACAAATTCCTGTACATCAACAGAG 2898
2883 TTCAGATTTCCCGTATCCCAACATTTTGAGAGAGTGCAGTGATCAAGTTGAGATGA 2942
2899 TAACTGGTTTCCCGTACCTTCGCACACTTTGCAAAATTTGGAGTTCGACGAGTTCGAACAA 2958
2943 GGAGGACACACTGAGATGCCAGGTGATGTGAATTTTTTTCCTACAATCGAAGTACGAAG 3002
2959 GAAGAACTACTCTTCGTGTGCGCTGATGTACACACTTCTTAATCAAAGGTATGAAG 3018
3003 GAGCGGTGACAAACAATTCCTCAATGTGTACACGATCGGTCTCATCTGAGATGATAGCGGTA 3062

Db 3019 GACACGTAATGTGCACGCTCTTCTGAAAGAAATCAGTTTCCAGGAAATGTTAGTGGG 3078
Qy 3063 AGGAGTACTAAACAGTGTCTTCCAAACACCTAAAGGGAATTTGTAACCTTCACTCAGG 3122
Db 3079 CTGGCTCTATCAATCTCTGTCTCAGCCGCTTAAGGGAATTTTGACCTTTCACACGT 3138
Qy 3123 CTGATAAATTTGAGTTAGAGGAGAAGGCTATAGAATGTGAACACCGCTTCATGAGATCC 3182
Db 3139 CTGACAGGAGGCCCTTCTCTCAAGGGCTATGCAGATGCCATCTGTACATGAGGTAC 3198
Qy 3183 AAGGAGAAACCTTTGAAGATGTGTGCTGTGTCAGATGACGGCACTCCACTGACTCTGA 3242
Db 3199 AAGGTGAGCGTATGACAGCGTATCGTTAGTTCGACTTAACACCTTACGCCCTGTATCTATCA 3258
Qy 3243 TTTCCAAAGTCTTCCCGCAGTCTTCTAGTCTGCTCTGACTAGACACACAAAGAGCTTCAAAAT 3302
Db 3259 TCGCAAGAGACAGTCCCGATGTTCTGTCTGTCTGTTGTCAAGACACACAAATCCCTAAAGT 3318
Qy 3303 ATTACACCGTAGTGTAGATCTCTTTAGTACAGATAATTTAGTGTCTTCTTTAAAGCT 3362
Db 3319 ACTACACCGTTGTGATGGATCTTTAGTTAGTATCATTTAGAGATTTAGAACGGGTTAGTA 3378
Qy 3363 CTTTCCCTTTAGAAATGTATGTAGACAGAGTAGTAGTAGCAATTTACAGATGATG 3422
Db 3379 GTTACTTTATAGACATGTACAAAGTAGATGACGACTCTCAATAGCAATTTACAGTCCACT 3438
Qy 3423 CAGTGTTCAAAGGTCATAATCTCTTCTGTGGCAACACCTAAATCAGGAGACTTTCCAGATC 3482
Db 3439 CTGTGTTTAAAMATTTCAATCTTTTGTAGCAGCTCCAAAGACTGGAGATATATCTGATA 3498
Qy 3483 TACAGTTCTATACGATGTATGCCCTCCCTGTTAATAGTACTATCTTAAACAGTATGATG 3542
Db 3499 TGCAATTTACTATGATAAGTGTCTTCTCGGAACAGCAGCTGTGTTGCAACAACTACGACG 3558
Qy 3543 CTGTTACATGAGGTTAGTGATTAATAGTCTTAATGTGAAGGATGTGTTCTTGATTTTT 3602
Db 3559 CTGTTACCATGAATTTGACTGACATTTCTCTGAATGTCAAGATTTGATATATGATATGT 3618
Qy 3603 CCAAAAGTATTCGATGCAAAAGGAGGTGAACCATCTGTAGAGCCAGTTTGGGTACCG 3662
Db 3619 CTAAGTCTGTAGCTGCTCCGGAAGATGTCAACCACTTTAATACCGATGGTACGAACGG 3678
Qy 3663 CGCGGGAACCGCCAAAGGCTGCAGGACTACTCGAAATCTGTTGGTGAATGATTAAGAA 3722
Db 3679 CGGCAGAAATGCTCGCCAGACTGCACTGTTGGAAATCTAGTTGCGATGATTAAGAA 3738
Qy 3723 ATTTCAAGCACACAGCTGACGGGGAGTGTGATGATGAGAGCACCGCATCTGTTCTAG 3782
Db 3739 ATTTTAATTCACAGAGTTGTCCGGAGTAGTTGATATTTGAAATACTGCAATCTTTAGTGG 3798
Qy 3783 TAGATAAGTTTTTTTGTAGCTATTTTATTAAGAAAGAAATATACACAAATAATTTGCTG 3842
Db 3799 TAGATAAGTTTTTTTGTAGTTATTTACTTAAGGAAAGAAAGAAACCAACAAATAATTTT 3858
Qy 3843 GAGTGATGACGAAGATTTCAATGATGAGATGGTTGGAAACAGAAAGAGTACTATTTGG 3902
Db 3859 CACTGTTTAGTAGAGTCTCTCAATAGTGGATAGCAAGCAAGAACAAAGTCAACAATTG 3918
Qy 3903 ACGACTTGGCTAACTACAAATTTTACAGATCTGCGGCCATCGATCAGTACAGACACATGA 3962
Db 3919 GTCAGTTGGCCGATTTTGTACTTTTGGGATCTTCAGCGCGTTGATCAGTACAGGCATATGA 3978
Qy 3963 TCAAGGCTCAACCAAAACAGAAATTTGACCTTTTCAATTCAGAAATGAAATACCTGCTGTC 4022
Db 3979 TTAAGCGCAACCGNAGCAGAACTGGATCTGTCATTTACAGACAGATATCCAGGTTGC 4038
Qy 4023 AAACAATTTGCTACATTCGAAGCAGATCAACGGTATTTTGG---CCGGTTTCTCAGAC 4079
Db 4039 AAACGATTTGTATCATTTCAAGAAATCAACGCAATATTTGGTCTCTTTTCACTGAGC 4098
Qy 4080 TTAACGTTTCTGCTCGAGCATTTGATTTCTAAGAGTTTCTTTTCTTTACTAGGAA 4139
Db 4099 TTCAAGGCAATTTACTTGACAGATTTGACTCAAGCAGATTTCTGTTCTTTACGAGAAAGA 4158

Qy 4140 CTCCAGAACAGATTCAGAAATTTTCTCGGATCTCGACTCGCAGCTTCTCTATCGATGTGT 4199
Db 4159 CACCGGCTCAGATCGAAGATTTCTTCGGAGATCTAGACAGTCTATGTCCTCAATGGACGTAC 4218
Qy 4200 TAGAACTGATATTTCTAAGTATGATAGTACAGAACGATTTCTATGCTGTAGAGT 4259
Db 4219 TTGAGTTGGATTTCTGAAGTATGACAGTCCCANAACGAGTTTCAATTTGCTGTGAGT 4278
Qy 4260 ATCAAAATATGAAAGATTTGGGTCTCAATGAGTTTGTGGCCGAAGTGTGAAACAAGGGC 4319
Db 4279 ACGAATCTGGAGAGACTGGGTCTGGAGGATTTCTTGGCAGAAGTGTGAAACNAGGGC 4338
Qy 4320 ACAGGAAAAACAATTTGAAGGATTAATTTGCTCGAATCAAGACATGTTCTGTGTATCAAA 4379
Db 4339 ATAGAAAAACCACTCTGAAAGATTAACACTGCTGTATAAAAACTGTTATGTGTACCAGA 4398
Qy 4380 GGAAGCGGTGATGTGACTACTTTTATCGGCAATACTGTTAATATAGCAGCTTGTCTGG 4439
Db 4399 GAAAGAGTGTGTATGTTTACAACTTTTATCGGTAAATACCGTCATCATTTCTGTTGCTTG 4458
Qy 4440 GTTCAATGTTTACCGATGGAAGGTCAATAAAGGTCTTTTGTGGAGACGATTTCCGTTT 4499
Db 4459 CATCAATGCTCCGATGGAAAAATTTGATAAAGAGCCTTCTCGGAGATGATAGTTGT 4518
Qy 4500 TGTATTTTCCAAAGGTTTGGATTTTCCCTGACATTCAGTCAATGTCATGCTCAATGTTGA 4559
Db 4519 TGTACTTTTCTAAGGTTTGTGATATCCCGATATACAAAGCTGCTTAATCTAATGTGA 4578
Qy 4560 ATTTGAGGCAAACTGTATAGAAAGGTACGGTTACTTTTGTGGTATGATACATCATAC 4619
Db 4579 ATTTGAGGCAAACTGTTCAAGAACCAATATCGGTACTTCTCGGAGGATGACGTGATC 4638
Qy 4620 ACCATGATAGGAGCAATAGTGTATTATGATCTTTGAAAGTTGATCTCCAAACTTCGGG 4679
Db 4639 ATCAGATAGAGTTGCAATAGTATATCTACAGCCCTTTGAAGCTGATTTTCGAACTTCGTG 4698
Qy 4680 CAAAAACATATCAAGGATTTATGATCACTTAGAAGATTTAAGGTGCTTTTGTGGATGTTG 4739
Db 4699 CTAACACATCAAGGATTTGGATCAATTTGGAGGATTTCAAGAGATCTCTCTGTGATGTTG 4758
Qy 4740 CTTGTTGCTCGGAACTGCTTAGGCTTTCCGAGCTGAAAGCAGCTATCAAGCAGG 4799
Db 4759 CTGAGTCTGT---GAACAATTCGCGATTAACACAAATTTGGAGCAGCTGTTGGGAGG 4815
Qy 4800 TTCAATAAACCGGATTTGATGCTGTTGCTTTTAAATTTGTTTAAACAAATTTTGTGTG 4859
Db 4816 TTCAATAAACCGCCCCACCTGTTGTTGTTTATAGAGTTTAGTTAAGTATTTGTCAG 4875
Qy 4860 ATAAATTTTATTTAGAACTTTGTTTTTAAATGGCTGTTAGTCTCAGAGATACGTCAAA 4919
Db 4876 ATAAAGTTTGTTTAGAAATTTATTTCTTGATGCTCTAGTTGTTTAAAGGTAAGGTAAT 4935
Qy 4920 ATTACGAGTTTCAATGATCTTTTGAACAGAGTATGATCTCCGCAATTCATGACTAAG 4979
Db 4936 ATTAATGAGTTTATCGATCTGCAAGCTCTGAGAACTTTTACCGTCAATGTTCCAGCT 4995
Qy 4980 GTCAAGAGTGTAGAAATATCGACTGTGCAAGAATTTAGGCTGTAAAGATGATGATGTT 5039
Db 4996 GTAAAGAGTGTATGTTTCAAGGTTGATTAAGATTTATGTTCCATGAAATGAATCATTTG 5055
Qy 5040 TCTGATGTAGATTTTCAATGAGGTTTAAAGTTAGTTAGTTAAGAAAGGTTATGTTGCTAGCT 5099
Db 5056 TCTGAAGTAAATCTCTTTAAAGGTTAAACCTTATAGAAGGTGGTATGTTTGTCTAGTT 5115
Qy 5100 GATTTGTTAGTGTCTGGGAGTGAATCTCCCGATTAACCTGCGTGTGTTGTCAGTGT 5159
Db 5116 GGTCTTGTGTGTCGGGTGAGTGAATTTACAGATTAATTTGCCGTGTTGTTGAGTGT 5175
Qy 5160 TGTATTTGTAGATAAGAAATGAAGAGGATGAAGAAAGCAACGCTGGGTGCTATCAAGCC 5219
Db 5176 TGCATGTTTACACAGAGATTTGAAGAGCGGACGACCACTGGGTCTATATTACACT 5235

5220 CTTGCTTGCAAAAGAAATTTCTTTTAAAGCTAAATCCCTAAATTAATCAATAACATCCGAG 5279
5236 GCTGCTGTAAGGCGGTTTCAGTTTAAAGTGGTCCCAATTAACGGTATTAACAAG 5295
5280 GATGCTGAGAAAGCACCGCTGGCAAGTGTAGTGAATATCAAGAGAGTGGCTATGGAAGAA 5339
5296 GATGCAGAAAAGAACATATGCGCAGGTCCTAGTAAATATTAATAATGTAATAATGAGTGG 5355
5340 GGATATGCTCTTTATCTTTGGAGTTCGTTTCAATTTGTGTAGTACATATAAATAATATGTA 5399
5356 GGCTACTGCTCTTTGTCATTAGAAATTTGTCTGTGTATTTGTTTATAAATAATATATA 5415
5400 AGAAAAGTTTGAAGGAACGTATTTGAGTGCACAGAGCGCTCGCAATTTGAATCACT 5459
5416 AATTGCGGTTGAGGAGAAAGTAACGAGTGTGACGATGAGGAGCCATGGAATTTCA 5475
5460 GAAAAGTTTGTGAGGAGTTCTGTGATGAAGTACCAATGCTGTGAACTCGAAAAGGTT 5519
5476 GAAGAAGTTTGTGATGAGTTCATGGAGAAATGTTCCAATGTCGTTAGACTCGCAAGTTT 5535
5520 CGGAAAACAAAAGAAATGGTAGTAATAATGTTAAATAT----- 5561
5536 CGAACCAATCTCTCAAAAAGAGTCCGAAAAATAATAATAATTTAGGTAAAGGCGCTTCA 5595
5562 ---AAGAAATAAATACAGTGTGTAAGAGGTTTAAATTCAGGAATTTGAGGTAAT 5618
5596 GCGGAAGGCTTAAACAAAAGGTTTGAAGTTGAAAGAGTTTGATTAATTTGAT 5655
5619 GTAAGTGTACGAGTCTATCGGCTCATCGAGTACGTTTAAATCAATATGCTTATACAA 5678
5656 GAAGATGAAGCGAGAGTTCGTCGGGATCTGATTCGTTAATATATGTTTACCCAA 5715
5679 TCAATCTCCGAGCCAAATTTGTTTACTTATCTTCGCTTACGAGATCTGTCAGCTGA 5738
5716 TCACTTCTCCATCGCAATTTGTTTGTGTCATCTGTATGGGCTGACCTATAGAAATGTT 5775
5739 TCATCTGTGTACAAATGATTTGGTAAACGAGTTTCAACGCAACAGCTAGGACAAAG 5798
5776 TAAAGCTTTGTACAAATCTGTTAGGTAAACGAGTTTCAACACAGCAAGCAAGAACTACTG 5835
5799 TCAACAGCAATTTGCGGATGCTGCAACCTGCTGCTAGTATGACAGTGAGATTTCTTG 5858
5836 TCAAAAGCAGTTACGAGGTTGGAACCTTTCCCTCAGACACCGTCAGATTTCTTG 5895
5859 CATCGGATTTCTATGTGTATAGATATAATTCGAGCTTGCATCGTTGATCAGCGGTTAT 5918
5896 GCGATGTTTATAAGGTGTACAGTACAATGCAATTTAGATCTTTTAAATTAATCTGCTGC 5955
5919 TAAATAGCTTTGATCTAGAAATAGAAATATAGAGGTTGATATCAACCGGACCGCAATA 5978
5956 TGGGGACTTTTCGATCTAGGAATAGCAATAATCGAAGTAGAAATCGGCAAGTCCGACAA 6015
5979 CTACTGAAATCGTTTAAACGAGCTCAGAGGTAGACGATGCTACTGTAGCTATAAGGGCTT 6038
6016 CAGCTGAAAGTTAGATGCTTACCGCAGAGGTAGACGCTACGTTGCAATTCGCTCTG 6075
6039 CAATCAATAATTTGGCTTAATGAATCGTTTGTGGAATCTGCAATGTTCAATCAAGCAGGT 6098
6076 CTATAAATAATTTAGTTTAACTAGTAAAGAGTACTGAGCTGTACATCAGAAATACTT 6135
6099 TTGAGATGCTAGTGAATGCTGACACCAACATCGGCTACTTACTAGCTATTTGTTGTAG 6158
6136 TTGAAAGTATGCTGCGGTTGCTGCGACCTCTGCACTGCAATTAATGCTAGTGTCT 6195
6159 ATTTCCTAAAATAAGTTCGCTGAAGACTTAAATATCAGGTTGCTGATACCAAAATCAG 6218
6196 GAAATATAAA-----GTTTGTGTTTCTAAACACACGTTGATGATGATAGCTAG 6247
6219 AGTGGTTGTTGCTCCACTTAAATATACGATTTGTCATATCTGATCCACAGCACTAAACCA 6278
6248 AGTGTGTTTCCCTCCACTTAAATCGAAGGTTAGTGTCTTGGAGCGCGCGGAGTAAACATA 6307
6279 TGTGATGGTGTATCTGTGTTATGCGGTAAACATCGGAGAGGTTTGAATCTCTCCCTAA 6338

Db 6308 TATGTT-TCATATATGTCGTAGGACGTAAAAAGCGGAGTTCGAATTTCCCGCGAA 6366
QY 6339 CCGCCGGTAGCGGCCA 6355
Db 6367 CCGCCGGTTGCGGCCA 6383

RESULT 13
TMO417701 6383 bp RNA linear VRL 31-OCT-2001
Tomato mosaic virus camellia strain, complete genome.
LOCUS AJ417701
DEFINITION AJ417701.1 GI:16580717
VERSION coat protein; complete genome; movement protein gene; RNA-depended
KEYWORDS RNA polymerase.
SOURCE Tomato mosaic virus
ORGANISM Tomato mosaic virus
VIRUSES; serNA positive-strand viruses, no DNA stage; Tobamovirus.
REFERENCE 1
AUTHORS Xue, C., Tao, X. and Zhou, X.
TITLE Molecular aspect of pathogenicity differences on Nicotiana tabacum
induced by Tobacco mosaic virus and Tomato mosaic virus
JOURNAL Unpublished
AUTHORS Zhou, X. and Xue, C.
TITLE Characterization and complete nucleotide sequence of two isolates
of tomato mosaic virus
JOURNAL Unpublished
AUTHORS Zhou, X. (bases 1 to 6383)
TITLE Direct Submission
JOURNAL Submitted (26-OCT-2001) Zhou X., Zhejiang University, Institute of
Biotechnology, Kaixuan Road 268, Hangzhou, Zhejiang, 310029, CHINA
FEATURES
1. 6383
/organism="Tomato mosaic virus"
/viral
/mol_type="genomic RNA"
/strain="camellia"
/db_xref="taxon:12253"
/country="China"
2. 4922
/note="contains readthrough stop codon"
/codon_start=1
/transl_except=(pos:3420..3422,aa:Gln)
/product="RNA-depended RNA polymerase"
/protein_id="CAD10425.1"
/db_xref="GI:16580718"
/db_xref="GOA:Q911R9"
/db_xref="TrEMBL:Q911R9"
/translation="MAYTQTATSSALLETVRGNNTLVNDLAKRRLYDTAVDFNARDR
RPNFNSKVVSEOTLIATKAYPEQITPTNTQNAVHSLAGLSLELEYLAWQIPYG
SITYDIGNFASHLPKRAYVHCMPNLDVRIMRHEGOKDSIELYLSLRGKHVP
NPKAPRYAEMPEVCHDTPTQCGHSECTGRVVALHLSYDIPADREGALL
RKNVCAAFHSENLLEDHSHNDINACFQDRDLTSPASESTLNTSHSVN
ILKYKTYTPASRNVKMEFLVRVNTWCFKSRIDTFLYKGVHAGVDSQEPYK
AMEDAWHYKTLAMCNSERILLEDSSVNYWPKMDVVIINGVTARSEWDVKSLQS
LVSRDPTVTVTHIRTYQALTYGNLVSFVESIRSRVINGVTARSEWDVKSLQS
LSMTFLTKLAVLKDILLISKFALPKTVSOHVDEISLAFGNAPPSIKERLNRKL
IKITENALIRVPDLYVTFFHRLVSEYKMSVMDPVLDIRKMEETEYNALSDLSVL
KNSDFVDVFSQMSLEVDPMATAKVIIVMSNESGLTLTFEQTEANVALALQDS
EKASGALVTSRDRVEELISIGSMARGELQAGLSDVPESYTRNEELSELEQFMA
TASGLIHKQMSIVYTGPKVQMKNFIDSLVASLSAASVNLVKILKDTAAIDLETRQ
KFGVLDAVKRWLPKPSKNAHWVTHARKYHVALLHDEFGIITCDNRVRVSS
ESVYSDMAKLTALKLKDGEHPVSSAKVLDVGVGCGKTEILSRVNFREDLTLV
PQQAEMIRRRANASGIIVATKDNRAVDSELMNYGKARCPKELFIDEGLMHTG
CVRLEWMSLCDIAVYGTQIIPYINRVTPPFAHFAKLEVEDETRTTLURCPAD
ATFLNQRVEGVMCTSEKMSQEMVSGAASINPVSKPLGELITFTQSKDEALLS
RGYADVTHVEVQGTADVSLVRLTPTPSIAGDSPHVLSLHSTKSLKYTYVM
DPLVSIIRDLERSVSYLLDMYKVDAGTQQLQVDSVFNKFNLPVAAKPTGIDIMQFY
YDKCLPSTLLNNDYATMKLTDIISLVKDCILDMSKSVAAKPKVPTLIIMVRTAA
EMPRQTGLLENLVAMIKRNFNSPELVGVVDIENTASLVVDKFDSDLSLLKKEKPKNF

SLFRESLNRIWIAKQEVTTIGQLADPFDVBLPAVDQVVRHMIKAPQKQKLDLSIQTEYP
ALQTVYHSKXKINAFGLFSELTRQLDLSIDSRFLFFTRKTPAQIETFFGLDSDHV
PMDVLIDSKYQNEFHCAYEYEWRLGLEDFLAENVKQKXRTLLKDTAGTIG
TCLWYQKSGDVTTFIGNVTIASCASMLPMELIKGAFCGDDSLYLPKGCYEPDI
QOANLWNPPEAKLKKQYGFPCGVYIHDRGCIYVYDPLKLSKLGAKHIDWHL
EEFRSLCDVAESLNNCAIYTLQDDAVEGHKTAPPGSFVKSLSVLYSLDKVLFRLS
LDGSSC"

CDS

72..3422

```

/codon_start=1
/product="RNA-dependent RNA polymerase"
/protein_id="CAD10426.1"
/db_xref="GI:16580719"
/db_xref="GOA:Q911R8"
/db_xref="TrEMBL:Q911R8"
/translation="MAVQTATSSALLLTVRGNNLTVNDLAKRLVYDAVDFNARD  
RPNKFSKVSSEETLLATKAYPEFOITFVNTQNAVHSLAGGLSLELYLRLERGNKHPV  
SLTYDIGNVFASHEEYVCHVCMENLVDYIMREHGOQDSIELYLSRLERGNKHPV  
NFQKADFAEMNEVYCHDTPQTCGHSQECYTGRTVYALHLSYIDIPADEFGALL  
RKNHVCYAFHSENLLEDSDHVNDEINACFORDGDRLTFFSASESTLNYSHSYN  
ILKVCYTFYPASNREYMKFELVTRVNTWFKFSRIDTFLLYKGVAKHGVDSQFYK  
AMEDAHYKKTLAMCNSERILLDESSVNVWFFPKMDWVLPVLPDISLETSKRTRKEV  
LVSRDFVYTVLHRTYOAKALTVSNVLSFVESIRSRVINGVTARSEVDVDSLLQS  
LSMTFFLHTKLAVLKDLLLSKALGPKTVSQHWDBISLAFGNAPSIKERLINRKL  
IKITENALEIRFDPVLYTFHRLVSEYKMSVDMPLDIRKMBETEEMYNALSDLSVL  
KNSDKFDVFSQCSLEVDPMTPAAKVIIVAVMSNESGLTLTFEQPTANVALAQDS  
EKASDGLAVTSDRDELSIKSGMARGELQAGLSDVPSSYTRNEEIESLQFHMA  
TASLIHKQMSIYVTGLKQVQMKNFIDSLVASLAAVSNLKLKIDTAIDLETRQ  
KGVLDVASKRWLKPASKHANGVETHARKYVALLHDEFGITCDNWRVAVSS  
ESVYVSDMAKRLTKLLKDGEPHVSAAKVLVDVPGCGKTEILSRVNEEDLLIV  
PGROAEMIRRRANASGIIIVATKDNVRAVDSFLMNYGKARQCFKRLIDEGMLHTG  
CVNPLVMSLCDIAYVGDYQOPIYINRVTPYPAHFAKLEVDFTRETRTLRCPAD  
ATHPLNYEGHVCTSEKRSYQEMVSGAASINPVSKPLGEIITFTQSDKEALLS  
RGVADVHTHEVOQETVADVSLVRLTPTVPSIIAGDSPHVLVLSLRHTSKLYTVVM  
DPLVSIIRDLERVSYYLLDMVKVDAGTQ"
```

CDS

4906..5700

```

/codon_start=1
/product="movement protein"
/protein_id="CAD10427.1"
/db_xref="GI:16580720"
/db_xref="GOA:Q911R7"
/db_xref="TrEMBL:Q911R7"
/translation="MALVVGKVNINEFIDLSKSKLLPSNFTPVKSVWVKVDKIMV  
HENLSEVLLKGLVIGGYVCLVGLVSGWNLVDNCRGGVSGVCMVDKRWERADE  
PTLGSYYTAAKFPQKVPVNTGITTDAKKNIVQVLNINKNVKSAGCPILSLFV  
SCVITYKNIKLGLREKTVSNDGPMELSEEVDEFEMVPMVSLRKLPRKTSKRG  
PKNNNLGKRGSGRPRPKPSDFEVGKEFDNLIDEAETSVDSDSY"
```

CDS

5703..6182

```

/codon_start=1
/product="coat protein"
/protein_id="CAD10428.1"
/db_xref="GI:16580721"
/db_xref="GOA:Q91WL6"
/db_xref="TrEMBL:Q91WL6"
/translation="MSVSIITSPQFVFLSSWADPIELNVCTNSLGNQFQTOQARTT  
VQOQFSWKPFPPOSTVFPDVKYRYNAVLDPLITALLGSPFDKRNRIIEVENQN  
PTTAETILDATRRVDDATVAIRSAINNVLNVLVRGTGLYNTFESMGLVMTSAPAS"
```

ORIGIN

```

Query Match 41.1%; Score 2612.6; DB 14; Length 6383;
Best Local Similarity 64.1%; Pred. No. 0;
Matches 4088; Conservative 0; Mismatches 2234; Indels 55; Gaps 8;

Qy 18 CGACAAACAACAAATTAACAAACATATTACAAACAAACAAACAAACAAATGGCAC 77
Db 23 CAACAACAACAACAACAACATTTACATTC-----TACAACATCAATGGCAT 78
Qy 78 ACATACATCTATTAATTAGCAACGCCCTTCTTGAACGGTGAAGTGTGTAACAACTCTCG 137
Db 79 ACACACAACAACGACACATCGTCCGCTTTGCTTGAGACCGTCCGAGGTAAACAATACCTTGG 138
Qy 138 TTAATGACCTTGAAGAGGCGATGTACGATCGGCGGTGGAGAGATTTAAAGCCCGCG 197
Db 139 TCAAGATCTTGAACAGGCGCTCTATATACACAGCGGTGATGAATTTAATGCTAGG 198
```

```

Qy 198 ACCGTAGAACCAAGGTCAACTTTTCAAAAATAATTAGCGAAGACAAACGCTTCTAGTCT 257
Db 199 ACCGAGGCTTAAGTCAATTTTCCAAAGTAGTAGAAGCAAGACGCTTATTGCAA 258
Qy 258 CCAACGGGTACCCGGAGTTCAGATTAACCTTTTATAACTCAAAATGCCGTACACAGTT 317
Db 259 CCAAGGCTACCCAGAAATCCAAATTTACATTTTCAACACGACGCAAGACGCTGTGCTCC 318
Qy 318 TGGCTGGAGTTTGGAGACATTAGAATTCGAATATCTGATGCTACAGTTCCCTATGGAT 377
Db 319 TTGAGCGGCTCTCCGATCATTAGAATTTGGAATATCTGATGATGCAAAATTTCCCTACGGAT 378
Qy 378 CGCCGACATATGATATAGTGGGAACCTTTGCAGCACATTTTGTTCAAAGCGAGGATTAACG 437
Db 379 CATTGACATATGATATCGGAGGTAAATTTTGCATCTCATCTGTTTCAAAAGCGGACATACG 438
Qy 438 TGCATTGCTGTATGCCCAATCTGGACATAGAGATATATAGAGGACGAGGACAAAGG 497
Db 439 TTCACTGCTGTATGCCGAATCTGGATGTCGCGACATAATGCGGACACGAGGGCCAAAGG 498
Qy 498 ACTCAATTGAGATGATTTGTCCAGATTTGCTCTTCTAAACAAGGTAATTTCCCTGAGTTTC 557
Db 499 ACAGTATAGAACTATACCTTTCTAGGCTCGAGAGGGGCAACAAACAGTCCCAACTTCC 558
Qy 558 AAAGGAGGCTTTTAAACAGGTATGCAGAACTCCCAACGAAGTCTGCTGCTCTTAAACCTT 617
Db 559 AAAAGGAAGCTTTTCGACAGATACGCTGAAATGCAACGAAGTAGTCTGTACAGATACCT 618
Qy 618 TTCAAGATTGTGCAATATACATCCGCCAGAGAAATAGTGTGTAAGATAGCTGTGCTCTGC 677
Db 619 TCCAAACGCTGTGGGCAATCTCAGGAAATGTTTACHACGGGAAAGAGTGTATGCTATTGCTTGC 678
Qy 678 ACAGTTTGTATGATATTTCTGTGTCATGAGTTTGGAGCTGCTGTTAAATCTTAAAGATATAC 737
Db 679 ATAGTATATACGATATACCTTGGCGAGGTTTGGCGGCGCACTGCTGAGAAAGATGTAC 738
Qy 738 ATGTATGTTATGACGCTTCCATTTTGGCAGAAGCATTTATCTAGACCAAGCGAGGTTA 797
Db 739 ATGTGTTTATGCGCGCTTTCCACTTTTCCGAGAATTTACTTCTCGAAGATTTACACGTC 798
Qy 798 CGTTTAAATGAATAGCGCAACTTTTCAAAAGAGAGAGTGTATGATGTTTCTTTTCTTTG 857
Db 799 ACCTCATGAGATTAATGATGCTTTTCAAAAGAGATGAGACAGGTTGACTTTTTTCTTTG 858
Qy 858 CTGATCAAAAGTACTTTAAATTTATAGTCATAAAATCAAAAATATCTTGCATTTATGTAGTTA 917
Db 859 CATCTGAGAGTACTCTTAATTTATAGTCATAGTTTATTCTAATATTCTTAAGTATGTTGCA 918
Qy 918 AATCTTTACTTTCTGCTTCTAGTAGAATAGTTTACTTTAAGGAATTTTGTAGTCACTAGGG 977
Db 919 AAACCTTACTTTCCAGCCTCTAATAGAGAGGTTTACATGAAGGAGTTTGTAGTAACATAGAG 978
Qy 978 TTAATACTTGGTTTGTAAATTTACCAAGTAGATACCTATATCTGTGTACAAAGAGTTTA 1037
Db 979 TAAATACCTGGTTTGTAAATTTTCTAGAAATAGATATCTTCTTTATTTGTACAAGGGTGTAG 1038
Qy 1038 GACAAAGTAGGCTGTAGTAGTATGATCTTCTATGAGCGGATGGAAGACGCTTTGCTTACA 1097
Db 1039 CGATTAAGGCTGTAGATAGTAGACGTTTAAAGGCTATGGAAGACGATGCACTACATA 1098
Qy 1098 AGAAAACTTTGGCCATGTTTCAACACTGAAAGAGCAATCTTTTAGAGACACGCGTTCGGTTA 1157
Db 1099 AAAAGACTCTTGGAGTGTGCAACAGTGAAGAAATCTTTGTAGAGGATTTCTTCATCAGTTA 1158
Qy 1158 ACTTTTGGTTCCCTAGATGAGGACATGCTGTGATAGTACCGCTGTTTGGAGGTTCTATT 1217
Db 1159 ATTACTGGTTTCCAAAAATGAGGGATATGCTGTATGTTTCACTATTGTGACATATCTCTCG 1218
Qy 1218 CCAGCAAAAAGATGACAAAGGAGTGAAGTCAATCTGTTAATCGTGACTTCGTTTACACAGTGC 1277
Db 1219 AGACTAGTAAAGAAACAGCAAGAGAGTCTTAGTTTCAGGGGACTTTGTTTATACAGTGT 1278
```

1278 TTAAATCATATCAGAAATATCAAGCCAAAGCGTTAACTTACCAGACGTTATTTCTTTCG 1337
1279 TAAATCACTTCGTACGTACAGCCAAAGCGTTAACTTACTCCAAAGCTTTATCTTTTCG 1338
1338 TGGAGTCTATAAGATCCCGCGTGATATCAATGCTGTACTGTAGTCTGGAATGGGATG 1397
1339 TCGAATCGATTCTGTCAGAGTGTATCAATTAACGGGGTTACTGTAGTCTGAGTGGATG 1398
1398 TAGATAAGCAATCTTTCAACCTTTGTCAATGATCTTTCTTCTTGGCAGACTAAAGTGGCTG 1457
1399 TCGATAAATCATTTATACAGTCTTGTCTGATGAGCTTTCTTCTTACATACCAAGCTTGCCG 1458
1458 CGCTTCAAGACGATATAGTAATGGGAAGTTTCGGTCTGGATATAGACCACCTTTCTGAC 1517
1459 TCTGAAAGACGATCTTTTGAATAGCAAGTTTGACCTGGTCCAAAACCTGTCTCAAC 1518
1518 TTATTTGGGATGAGTGGGCAAAATTTTTTGGAAACGTTTTTCCCACTATCAAAAGAGAT 1577
1519 ATGTGTGGGATGAGATTCTCCCTAGCTTTCGGCAATGCTTTCCCATCGATCAAGGAAAT 1578
1578 TGGTAGCAGGAAATTTCTGGATGTAGTGAGAAATGCTCTGAAAGATCAAGATCCAGATC 1637
1579 TGATAAACCGGAACTGATCAAAATTTACGGAGAAATGCGTTAGAGATCAGGGTCCCGATC 1638
1638 TGTATGTCACATGGAAGACAGGTTCTGTAGCTGAATACCAAGTCTTGAGGAGTTACCGC 1697
1639 TTTATGTCTCTTCCATGATAGGTTAGTTTCTGAGTACAAATGTCTAGTGGACATGCCGG 1698
1698 ATCTAGATATCAAGAGGACTTAGAAGAGCTGAGCAAAATGTACGACGCGTTATCAGAAAT 1757
1699 TGTAGACATTAGGAAGAGATGGAAGAACTTGAAGAAATGTACAAATGCACTGCTCCGATC 1758
1758 TATCTATCTTAAAGGCTGTGATAATTTTCAATATCGGAAAGTTTCAAGACATGTGCAAGG 1817
1759 TGTCTGTACTTAAATAATTCAGACAAGTTTCGATGTGATGTTTCTCCAGATGTGCCAGT 1818
1818 CTTTAGATGTAGTCTCTGATGTGGCAGCAGTATCTGTCAGTGGCCGAGANTAGAA 1877
1819 CTTTAGAAGTCGATCCCAATGACTGCAAGCAAGGTAATAGTAGCAGTTATGAGCAACGAGA 1878
1878 GCGGTTTAACTCTTACTTTTGAAGCAACCGAGGAGATGTGGCTTAAGGCTCTTAAA - 1936
1879 GTGGTCTTACTCTACGTTTGAACAGCCACCGAGGCTAATGTGCGCTAGCATGCGAAG 1938
1937 -----AGCACGGCTCTGAGCCGCTGTGTATGTCTTGAAACCGACATCCGAAGAGTGAAAG 1991
1939 ATTCTGAAAGGCTTCTGATGGGCGTTGGTAGTTACCTCAAGAGATGTTGAGCAACTGT 1998
1992 TAAATAAATTTCTATTGCTGAGAAAGGAGATGCTGTGTGTGCAAGAAAGTCATGGTT 2051
1999 CCATAAAGGCTTCGATGGCCCGTGGTGTGAGTTGCAATTTGGCCGATATCTGGCGAGCTTC 2058
2052 TGACGAATGCTAACTTAGACACACAGAGTTGGAGTCCCTCAACGATTTCCATAAGGCTT 2111
2059 CTGAACTTCAACACTAGAAACGAGAGATTTGAGTCTCTGAGCAGTTTCATATGGCAA 2118
2112 GCGTGGATGTGATTTAAAGCAAAATGCGATCGGTTGTCTACACTGGCTCTACTCAAAG 2171
2119 CAGCTAGTTCTGTTAAATCAATAGCAGATGTGTTCCGATCGGTGTACACGGGCGCTCTTAAAG 2178
2172 TTCAACAAATGAAGACTATGTGGACAGTTTGGCAGCTTGGTGTGCGCCACTGTATCAA 2231
2179 TTCAACAAATGAAGAACTTTATAGACAGCTGTGTAGCTTCGCTCTCTGCTGCGGCTGTGCA 2238
2232 ATCTATGCAAGTCACTAAAGGATGAAGTCGGGTATGATTTCTGATTCAGGGGAGAAAGTTG 2291
2239 ATCTAGTGAAGATCTTAAAGATACAGCCGGATTTGACCTTGAATCTCGTCAAAAGTTTCG 2298
2292 GTGTTGGGATGTCACTTTGAAAAGTGGCTCTCAAACCTCGCGCCAAAGGTCATTCAT 2351
2299 GAGTCTCTGGATGTGCTTCGAAAAGTGGCTAGTTAAACCATCGCAAAAGCAACATGCAT 2358
2352 GGGAGTGTCTGGATTAACAGGGGAAATGTTTACTGCACTTCTATCTTATGAAGGAG 2411

2359 GGGGGGTTGTTGAGACTCATGCGAGGAATATACGCTCGCATTTACTTGGACACGATGAAT 2418
2412 ATAGAAATGGTGTACTGAGACGACTGGAGGAGGTGGCTGTATCATCTGTATACAAATGGTAT 2471
2419 TTGGCATTTATCACGTCGGATAACTGCGCAGCGGTGGCTGTGAGTCTGTAGTCTGGTAGTAT 2478
2472 ATTCTGATATTGCAAGAGCTCCAAAATCTGAGGAAACAAATGAGAGACGGTGAACCCACG 2531
2479 ATTCTGATATGGCTAAACTCAGGACTCTGAAAAGATTTGCTTAAAGATGGAGAACCAACG 2538
2532 AACCTACTGCAAGATGTGACTTGTGATGGGGTGGCTGTGTTGTGAAAGTACAAAGGAG 2591
2539 TTAGTTTCAAGAAAGTGGTTTTGTTGATGGGCTTCCAGGGTCCGGGAAACAAAGGAAA 2598
2592 ATTTTGAAGAATTTGATCTTTGATGAGGATTTGATCTTTGTTCTGTGAAAACAAAGTGGCTG 2651
2599 TTTCTTTCGAGAGTTAATTTTGAAGAAGATCTAATTTCTGTCCCTGGTCTGCTCAAGCTGCCG 2658
2652 CTATGATCAGAGAGGCTTAATTTCACTGTGGACTGATAGAGCCCAATGCAANTGTGA 2711
2659 AGATGATCAGAGAGAGCTTAATGCGTCCGGCATTAATAGTGGCTCAAAAGGATAATGTGC 2718
2712 GAACGGTAGATTCACTTTCTAATGCAATCCAAAACCCGCGATCAC-----ACAAGAGGC 2762
2719 GCGCCGTCAATTCATTTCTTGTGTAATTAACGGGAAAGGGCACGCTGTCAATCAAAGAT 2778
2763 TTTTATTTGATGAAGGTTGATGCTGCACACCGGTTGTGTAACTTCTTGGTCTTATCT 2822
2779 TGTTCATAGACGAAGTTTGTGATGCTGCATCTGTTGTGTGTAATTTCTTGGTTGAAATGT 2838
2823 CTGGTTGCCACATCGCATACATTTACGGAGATACACAGCAGATTCCTTTTCATTAACAGAG 2882
2839 CTCGTGCGATATGTGATATGTTTATGAGACACCAACAAATTCCTGTACATCAACAGAG 2898
2883 TTCAGAAATTTCCGCTATCCCAACATTTTGGAGAAGCTGCAAGTGGATGAAGTTGAGATGA 2942
2899 TAACTGGTTTCCGCTACCTTGACACTTTTGCAAAATTTGAGGTCGAGAAAGTCGAAACAA 2958
2943 GAGGACCACTGAGATGCCAGGTGATGTGAAATTTTTTCTCAATCGAAGTACGAAG 3002
2959 GAAGAACTACTCTTCTGTTGCGGCTGATGCCACACATTTCTTAAATCAAAGGTATGAAG 3018
3003 GAGCGGTGACCAACCTTCAACTGTATACAGATCGGTCTCATCTGAGATGATAGGCGGTA 3062
3019 GACACGTAATGTGACGCTCTTCTGAAAAGAAATCAGTTTCCAGGAAATGGTTAGTGGGG 3078
3063 AGGAGTACTTAAACAGTGTTCCAAACCACTAAAGGGAAATTTGTAACCTTTCACTCAGG 3122
3079 CTGCGTCTATCAATCTCTGTGTCGACCGCTTAAAGGGAAATTTTGACCTTTCAACAGT 3138
3123 CTGATAAATTTTGAAGTTAGAGGAGAGGCTTAAAGAAATGGAACACCGTTTCATGAGATCC 3182
3139 CTGACAAAGGAGGCCCTTCTCTCAAGGGCTATGACAGATGTCCATCTGTACATGAGGTAC 3198
3183 AGGAGAAACCTTTGAAGATGTGCTCGTGTGATGAGTTGACGGCACTCCACTGACTCTGA 3242
3199 AAGGTGAGACTTATGACAGCGTATCGTTAGTTTCCACTAAACACCTACGCTGTATCTATCA 3258
3243 TTTTCAAGTCTTCCCGCATGTTTCTAGTCTGCTCTGACTAGACACACAAAGAGCTTCAAAAT 3302
3259 TCGAGAGACAGTCCGCAATGTTCTGCTGTGTTGTCAGACACACAAATATCCCTTAAAGT 3318
3303 ATTACACCGTGTGTTAGATCTTTTATGACAGATAATAGTGTGTTCTCTTCTTTAAGCT 3362
3319 ACTACACCGTGTGATGGATCTTTTATGTTAGTATCATAGAGATTTAGAACGGGTAGTA 3378
3363 CTTCTCTTTTGAAGATGTATATGTTAGAGACAGGTAGTAGATAGCAATTTACAGATGATG 3422
3379 GTTACTTATTTAGACATGTACAAAGTAGTGCAGGTACTCAATAGCAATTTACAGTGCAGT 3438
3423 CAGTGTCAAAGGTCAATATCTTTTGTGGCAACACTTAATCAGGAGACTTTTCCAGATC 3482

Db 3439 CTGTGTTTAAATAATTTCAATCTTTTGTAGCAGCTCCAAAGACTGGAGATATATCTGATA 3498
QY 3483 TACAGTTCTATTACGATGATGCTCCCTGGTAAATAGTACTATATCTTAACAAGTATGATG 3542
Db 3499 TGCATTTTCTATGATAAGTGTCTTCCGGGAACAGACGTTGTTGAACAACTACGACG 3558
QY 3543 CTGTACCAATGAGGTTTACGTGATAATAGTCTTTAAATGTGAAGGATTTGTCTTGTGATTTT 3602
Db 3559 CTGTTACCAATGAAATTTGACTGACATTTCTCTGAATGTCRAAGATTCATATTAGATATGT 3618
QY 3603 CCAAAAGTATTCGGATGCCAAAGGAGGTGAACCAATGTCTAGAGCCAGTTTTCGTACCG 3662
Db 3619 CTAAGTCTGTAGTCTGCTCGAAGAGTGTCAAACCAACTTTAATACCGATGCTACGAACGG 3678
QY 3663 CGCGCGAACCAGGGCTGCAGACTACTCGAAAACTGCTGTCGAATGATTTAAAGAA 3722
Db 3679 CGCGAGAAATGCTCGCCAGACTGACCTGTTGGAAAACTTAGTTCGGATGATTTAAAGAA 3738
QY 3723 ATTTCAACGACACAGACCTGACGGGACGATTTGACATTTGAGAGCACCGCATCTGTTGTAG 3782
Db 3739 ATTTTAAATTCACAGAGTGTCCGAGTAGTTGATATTGAAAAATACTGCATCTTTAGTGG 3798
QY 3783 TAGATAAGTTTTCATAGCTATTTTATTTAAAGAAATATACACAAAAATATTGCTG 3842
Db 3799 TAGATAAGTTTTCATAGTATTATTACTTAAAGAAATAAGAAAAACCAACAAAAATTTT 3858
QY 3843 GAGTGATGACGAAGGATTTCAATGATGAGATGGTTGAAAAACAGGAAGAACTACTATTGG 3902
Db 3859 CACTGTTTAGTAGAGAGTCTCTCAATAGTGTGATGACAAAGCAAGCAAGTCAATTTG 3918
QY 3903 AGCATTTGGCTTAACTACAATTTTACAGATCTGCCGGCCATCGATCAGPACAAAGCACATGA 3962
Db 3919 CTCAGTTGGCCGATTTTGATTTTGTGGATCTTCAGCGCGTTGATCAGTACAGGCATATGA 3978
QY 3963 TCAAGGCTCAACCAAAACAGAAATTTGACCTTTCAATTTCAAGTAAATACCTGCTGTC 4022
Db 3979 TTAAAGCCAAACGAGCAAGAACTGGATCTGTCATTTTCAAGACAGATATCCAGCGTTGC 4038
QY 4023 AAACAATTTGTCTACATTTCSAAGCAGATCAACGGTATTTTGG---CCGGTTTCTCAGAGC 4079
Db 4039 AAACGATTTGTATCATTTCAAGAAATCAACGCAATATTTGGTCTCTTTTCAGTGNAC 4098
QY 4080 TTCAAGGTTGCTGCTCAGGCAATTTGATTTCTAAGAAGTTTCTTTTCTTACTAGGAAAA 4139
Db 4099 TTCAAGGCAATTTACTTGACAGTATTTGACTCAGCAGATTTCTGTTCTTTACGAGAAGA 4158
QY 4140 CTCAGAAACAGATTTCAAGAAATTTTCTCGGATCTCGACTCGACGTTTCTTATGATGTTGT 4199
Db 4159 CACCGCTCAGATCGAAGATTTCTTCGAGATCTAGACAGTCAATGTCCTAATGGACGTAC 4218
QY 4200 TAGAACTGGATTTTCTAAGTATGATAGTCAAGAACGAGTTTTCATTTGTGCTGTAGAT 4259
Db 4219 TTGATTTGATTTTTCGAAGTATGATAAGTCTCAAAACGAGTTTCAATTTGTGCTTTCAGT 4278
QY 4260 ATGAAATATGAAAGATTTGGTCTCAATGATTTTGGCGAAGTGTGGAACAAGGGC 4319
Db 4279 ACGAATCTGAGGAGACTAGGCTGAGGATTTTCTTGGCAGAAAGTGTGGAACAAGGGC 4338
QY 4320 ACAGAAACAACTTTGAAGGATTTACATTTGCTGGAATCAAGACATGTTCTGTGGTATCAAA 4379
Db 4339 ATAGAAAAACCACTCTGAAGATTTACATGCTGCTGATATAAAACGTTTATTTGTGACAGA 4398
QY 4380 GGAAGCGGTGATGTGACTATTTTCAATCGGCAATCTGTTTATTAATAGACAGTTCGTGG 4439
Db 4399 GAAAGAGTGGTATTTTACAACTTTTATCCGTAATACCGTCATCATTTGCTGCTGCTTG 4458
QY 4440 GTTCAATTTTACCGATGGAAGGTCATAAAGTGTCTTTTGTGGAGAGATTTCCGTTT 4499
Db 4459 CATCAATGCTCCCATGGAATAATTTGATTAAGGAGCTTCTCGCGAGATGACAGTTTGT 4518
QY 4500 TGTATTTTCCAAAGGTTTGGATTTCCCTGACATTTCACTGATGCTGCTTAATCTCATGTGGA 4559
Db 4519 TGTACTTTTCTTGAAGGTTGTGAGTATCCCGATATACCAACAGGCTGCTTAATCTAATGTGA 4578

QY 4560 ATTTTGAGGCCAAACTGTATAGAAAGAGGTACGGTACTTTTGTGGTAGATACATCATAC 4619
Db 4579 ATTTTGAGGCCAAACTGTTCAAGAAAGCAATATGGGTACTTTCTCGGGAGGTACGTGATTC 4638
QY 4620 ACCATGATAAGGGAGCAATAGTGTATTATGATCTTTTGAAGTTTGTATCTCCAAACTTCGGG 4679
Db 4639 ATCAAGATAGAGTTGATATAGTATATCAACGCCCTTTGAAGCTGATTTTCGAACTTCGGTG 4698
QY 4680 CAAACAATATCAAGGATTTATGATCATCTTGAAGAGTTTAAAGGTTGCTTTTGTGCGATGTTG 4739
Db 4699 CTAACACATCAAGGATTTGGATCAATTTGGAGGATTCAGAAGATCCCTCTGTGATGTTG 4758
QY 4740 CTTGTTCCGCTCGGAACCTGGTCTTAGGCTTTCCGAGCTGGAACGAGCTATCAAGGAGG 4799
Db 4759 CTGAGTCGTT---GAACAATTCGCGATTTACACACAATTTGGACGAGCTGTTGGGAGG 4815
QY 4800 TTCTAAAAACCGGATTTGATGCTTTCTTTTAAATTTGTTTAACTTTTGTGTTG 4859
Db 4816 TTCTAAAAACCGCCACCTGGTTGTTTGTATTAAGAGTTTAGTTAAGTATTTGTCAG 4875
QY 4860 ATAAATTTTATTATAGAACTTTGTTTAAATGGCTGTTAGTCTCAGAGATCTGTCAAA 4919
Db 4876 ATAAAGTTTGTTTAGAAAGTTTATTTCTTGATGGCTCTAGTTGTTTAAAGGTAAGGTAAT 4935
QY 4920 ATTAGCGAGTTCAATGATCTTTCCGAACAGGATGAGATCTTCCGGCATTTCATGACTAAG 4979
Db 4936 ATTAATGAGTTTATCGATCTGTCAAAAGCTTGAGAAACTTCTCCGCTCGATGTTTCAGCCT 4995
QY 4980 GTCAAGAGTTTGAATATCGACTGTGGACAAGATTTAGCTTTAAGAAATGATAGTCTT 5039
Db 4996 GTAAAGAGTTTATGTTTCAAGAGTTGATTAAGATTTATGTTCCATGAAATGAATCATTTG 5055
QY 5040 TCTGATGTAGATTTACTTAAGGTGTTAAGTTAGTTAAGAAAGGTTAGTGTGCTTAGCT 5099
Db 5056 TCTGAAGTAAATCTCTTAAAGGTGTAAACCTTATAGAAGTGGATATGTTTGTCTTAGTT 5115
QY 5100 GATTTGGTAGTGTCTCGGGAGTGGAAATCTCCGGATAACTCCCGTGTGCTGTCAGTGT 5159
Db 5116 GGTCTTGTGTTGTCGGTGAGTGGATTTTACAGATTAATTTCCGCTGGTGTGAGTGTC 5175
QY 5160 TGTATTGTAGATAAGAAATGAAGAGTAGTAAGAAACAACTGCGGTGCTATCAAGCC 5219
Db 5176 TGCATGTTTACAAAGAGATTTGGAAGCGGACCAACCCACACTGGGGTCAATATTACACT 5235
QY 5220 CTTGCTTGCAAAAGAAATTTTCTTTAAAGTAAATCCCTAAATTTCAATTAACATCCGAG 5279
Db 5236 GCTGCTGTAAAGCGGTTTCAAGTTTAAAGTGGTCCCAATTTACGGTATTACAAAG 5295
QY 5280 GATGCTGAGAAGCACCCGTGGCAAGTGTGTAATATCAAAAGGAGTGGCTATGGAAGAA 5339
Db 5296 GATGCAGAAAAAGAACATATGGCAGGTTTAGTAAATATTAATAATGTAATAATGAGTGG 5355
QY 5340 GGNATCTGCTTTATCTTTGGAGTTCGTTTCAATTTGTGTAGTACATAAAAAATATGTA 5399
Db 5356 GGTACTGCTCCCTTTGTCATTAGAAATTTGTCTGTGTTATTTTATATAAAATAATA 5415
QY 5400 AGAAAAAGTTTGAGGGAACGTTATTTTGTAGTGTGACAGCGCTCGCAATTTGAACCTCACT 5459
Db 5416 AAATTTGGTTTGGGGAGAAAGTAACGAGTGTGAACGATGGAGGCCATGGAACCTTCA 5475
QY 5460 GAAAAGTTTGTGAGGAGTTCTGGATGAAGTACCAATGGCTGTGAACTCGAAAAGGTT 5519
Db 5476 GAAAGAGTTTGTGATGAGTTTCATGGAGATTTTCAATGTCGTTAGACTCGCAAGTTT 5535
QY 5520 CCGGAAAACAAAAAAGAAATGTTAGTAAATATGTTAAAT----- 5561
Db 5536 CGAACCAATCTCTAAAAAGAGGTCGAAAAATAATAATTTAGGTAAAGGGCGTTCA 5595
QY 5562 ---AAGAAATATAACAGTGTGAAGAGGCTTTTAAATTTGAGAAATTTGAGGATTAAT 5618
Db 5596 GCGGAAGGCTTAAACCAAAAGTTTGTGATGAAGTTGGAAAGAGTTTGTATTAATTTGAT 5655

```
QY 5619 GTAAGTATGACGAGTCTATCGCGTCAATCGAGTACGTTTAAATCAATATGCTTATACAA 5678
Db 5656 GAAGATGAAGCCGAGAGCTCGGTCGGGATCTGATTCGATTAAATATATGCTTACTCAA 5715
QY 5679 TCAACTCTCCGAGCAATTTGTTTACTTATCTCCGCTTACGAGATCCGTGTCAGCTGA 5738
Db 5716 TCACTTCTCATCGCAATTTGTTTGTGTCATCTGATGGCTGACCTATAGATTGT 5775
QY 5739 TCAATCTGTGTACAAATGCAATCGGTAAACGAGTTTCAACGCAACGAGCTAGGACACAG 5798
Db 5776 TAAACGTTTGTACAAATTCGTTAGTAAACGAGTTTCAACGCAACGAGCAAGCAACTACTG 5835
QY 5799 TCAACGAGCAATTTGCGATGCTCGAAGCTGCTGCTAGTATGACAGTACGATTTCCCTG 5858
Db 5836 TCAACGAGCAATTTGCGAGGCTGGAAGACCTTTCCCTCAGAGCAGCTGAGATTTCCCTG 5895
QY 5859 CATCGAATTTCTATGTGTATAGATATAATTTGCGAGCTTGTATCCGTTGATCAACGCGCTTAT 5918
Db 5896 GCGATGTTTATAGGTGTACAGGTACATCGAGTTTAGATCCTCTAATTTACTGCGTTGC 5955
QY 5919 TAAATAGCTTTGTATACAGAAATAGAAATAGAGTTTGAATATCAACCCGCAACGGAATA 5978
Db 5956 TGGGCTCTTTTGTATCTAGCAATAGAAATAGAAATAGAAATAGAAATAGAAATAGAAATAG 6015
QY 5979 CTACTGAATTCGTTAAGCGCACTCAGAGGCTAGACGATGCTAGTATGAGGCTT 6038
Db 6016 CAGCTGAACGTTAGATGCTACCGGAGGCTAGACGATGCTAGGTTGCAATTCGATCTG 6075
QY 6039 CAATCAATATTTGGCTTAATGAATGCTGCTGGAACCTGCAATGCTTCAATCAAGCAGGCT 6098
Db 6076 CTATCAATATTTAGTTTAAATGAATAGTAGAGTACTGAGCTATACATCAAAATACCT 6135
QY 6099 TTGAGACTGCTAGTGGACTTGTCTGGAACCAACTCCGGCTACTTACGTTTGTGAG 6158
Db 6136 TTGAAAGTATGCTGGGTTGCTGCGACCTCTGCACCTGCATCTTAAATGCATAGGTGCT 6195
QY 6159 ATTTCTCAATTAAGTCGCTGAGACTTAAATTCAGGCTGCTGATACCAATCAGC 6218
Db 6196 GAAATATAAA-----GTTTGTGTTTCTAAACACACAGCTGCTAGTACGATTAACGTAC 6247
QY 6219 AGTGGTTGTTGCTCCACTTAAATATACGATTTGCTATCATCTGATCCCAACAGTTTAAACCA 6278
Db 6248 AGTGTGTTTCCCTCCACTTAAATTCGAAGGTTAGTGTCTTGAGGGCGCGAGTAAACATA 6307
QY 6279 TGTGATGGTGTACTGTTGATGGCTGATGCGTAAACATCGGAGAGGTTGGAATCTCCCTAA 6338
Db 6308 TATGGT-TCATATATGTCGCTAGCAGTAAAGGAGGAGGATTCGAAATTTCCCGCGAA 6366
QY 6339 CCGCGGCTAGCGGCCA 6355
Db 6367 CCGCGGCTAGCGGCCA 6383
```

RESULT 14

```
AB069853
LOCUS
DEFINITION
Pepper mild mottle virus gene for 126 kDa protein, linear
ACCESSION
AB069853
VERSION
AB069853.1 GI:18652912
```

```
KEYWORDS
Pepper mild mottle virus
SOURCE
Pepper mild mottle virus
ORGANISM
Viruses; serRNA positive-strand viruses, no DNA stage; Tobamovirus.
```

REFERENCE

```
1
Hagiwara, K., Ichiki, T. U., Ogawa, Y., Omura, T. and Teuda, S.
A single amino acid substitution in 126-kDa protein of Pepper mild
mottle virus associates with symptom attenuation in pepper; the
complete nucleotide sequence of an attenuated strain, C-1421
Arch. Virol. (2002) In press
2
Hagiwara, K., Ichiki, T., Ogawa, Y., Omura, T. and Teuda, S.
Direct Submission
TITLE
Submitted (10-AUG-2001) Kyoji Hagiwara, National Agricultural
```

Research Center, Laboratory of Virology; Kannondai, Teukuba, Ibaraki
305-8666, Japan (E-mail: kyoji@affrc.go.jp, Tel: 81-298-38-7030,
Fax: 81-298-38-8929)

FEATURES

```
Location/Qualifiers
source
1..6357
/organism="Pepper mild mottle virus"
/mol_type="genomic RNA"
/isolate="C-1421"
/db_xref="taxon:12239"
70..3423
/codon_start=1
/product="126 kDa protein"
/protein_id="BAB84693.1"
/db_xref="GI:18652913"
/translation="MAYTQATNMAALASTLRGNPLVNDLANRRLYESAVEQCNADHR
RKVNPLRSISEEQTLIATKAYPEFOITFYNTQNAVHSLAGLRSLSELEYLMQIPY
STTYDIGNFAAFMRKGRDYVHCMPNDLRDMRHNAAQKDSIELYSLKMKQKVI
PYKPCFQDTPDQSPQVCGSPKFCHEGSHCTDKVYVALHSLYDI PADEFGAALIR
RNHVYCAAFHSENLLEDSDYSLDDIGAFPSREGDMLNFSVAESTLNYTHSYSNV
LKYCKTYPASSRREYMEFLVTRVWFCKESRLDTFVLGVGVHVRGVEQFYSKA
MEDAHYKKTLLAMNSERILLEDSSVNTWFFPMKDMVIVPLFDVSLQNEGKKLARKE
VWVSKDFVYTVLNRHTYQSKALTANVLSFVESIRSVIINGVTARSVDWDKALLQ
SLSMTFFLQTLAKMLKDLVQKFSKSLTEYVWDEITAAAFHNCFTPKERLNNK
LITVSEKALIKVPLDYTFHDLRVKXSVSEMPVLVDKLSLEAEAVMYNALSLSI
LXDSKEDVDVFSRMCTLGVDPLVAAKTVAVVSNESGLTLTFEPRTANVALQSP
TITSEBGSLSKIVSSDVGBESSIKEAVRKSEISHLGTLGNTVDSDEFORSEIESLQPH
MYSTETIRKQHMVYVYTGPKVQCKNYLDSLSAASVSNLKKI IKDTAIDLDET
KEKGFYDVLKLLKPLSGHGWVNDSDYKCFVALLTYDGENIYVCGEWRVVAV
SSESIVSDMGKIRAIRSVLKGEPHISAKVTLDVGVPCGKTKRILSRVNFEDIV
LVPGQAAEMIRRRANSGLIVATKENVTVDSFLMNYGRGPCQYKRLFLDEGLMLHP
GCYNELVNSLCSSEAFVGTQOIPVINEVATFPVPKHLSOLEVDVAVETRTTLCRPA
DITFLNQKTEQVWMCSTTSVRSVSEHVIQGAAMVNPVSKPLGKVIITFTQSDKSLLL
SRGVEDVHTVHEVQGETFEDVSLRLTPTPVGLIISKQSPHLLVLSLRHRSIKYITVV
LDVAVSVLRDLKCVSYLLDMYKVDVSTQ"
```

ORIGIN

```
Query Match 39.8%; Score 2529.6; DB 14; Length 6357;
Best Local Similarity 63.7%; Pred. No. 0;
Matches 4063; Conservative 0; Mismatches 2259; Indels 58; Gaps 12;
QY 6 TTTAATAGTTTTCGACAAACAACAATTAATAACAAAAACAATATTACAAACAACAACA 65
Db 6 TTTTTCACATTTAACACAACAACAACAACAACAACAACAACAACAACAACAACAACA 64
QY 66 CAACAATGGCACACATACATCTTATAATTAGCAACGCCCTTCTTGAAGAGCGTGAGTGA 125
Db 65 CTACAATGGCTTACACACAACAAGCTACCAACGCCCATTTAGCAAGTACTCTCCGAGGA 124
QY 126 AAAAACTCTCGTTAATGACCTTGCAGAGAGCGCATGTACGATACGGCCGTGAGAGAAT 185
Db 125 ATAAACCCCTTGGTGAACGACTTTTGTCTAATCGGAGACTGTACGAATCAGCGGTGCAACA 184
QY 186 TTAACGCCCGCGACCTAGACCAACAGCTCAACTTTTCCAAAACTATTAGCGAAGAGCAAA 245
Db 185 GCATGCAATGACCCGACGCCCAAGCTTAATTTTAAAGTTCGATAAGCGAGAGCAGA 244
QY 246 CGCTTCTAGTCTCCACAGCGTACCCGAGGTTCCAGATTACCTTTTATATACTCAAAATG 305
Db 245 CGCTTATCGCAACTAAGGCTACCTCTAGTTCCAAATCAGCTTCTACACAACGAGAACG 304
QY 306 CGGTACACAGTTGGCTGGAGGTTTGAGAGCATTAGAAATGGGAATATCTCATGCTACAAG 365
Db 305 CTGTGCACAGTCTCGCAGGTGGACTTCGGTCTTTTGGAACTAGAAATATCTTATGATGAGA 364
QY 366 TTCCTTATGATCGCGACACATATGATATAGTGGGAACTTTGAGAGCAGCATTTGTTCAAAG 425
Db 365 TCCCTTACGGTTCAACGACATATGATATCGGGGAAATTTTGTCTGCATCATGTTTAAAG 424
QY 426 GCAGGAGTATCAGTGCATTTGCTGTATGCCCAATCTGGACATACGAGATATATGAGGACG 485
Db 425 GTCGTGACTAGTTTCATTGCTGTCATGCCTAACATGACCTTACGTGACGTATCGCTCACA 484
QY 486 AAGGACAAAGGAGTCAATTTGAGATGATTTTGTCCAGATGTTCTCGTCTTCTTAACAGGTAA 545
```

Db 485 ATGCTCAAAAGGATGACATTTGAACCTGTACCTTTCAAGCTTCGGCAAAAGAAAGGTAA 544
Qy 546 TTCTCGAGTTTCAAGGGAGGCTTTTAAACAGGTATGCGAAGCTCCCAACGAAGTCTGCT 605
Db 545 TACCGGCATATCAAAAGCCATGCTTTGATAAATACACGAGCATCCGCAATCAGTAGTGT 604
Qy 606 GCTCTAAAACCTTTTCAGGATTTGCGAATACATCCGCCAGAGAAATAGTGTAGAGATACG 665
Db 605 GCTCGAAACCTTTTCAGCACTCGGAAGGGCTTCGC---ACTGCACGGATAAAGTATACG 661
Qy 666 CTGTTGCTCTGCACAGTTTGTATGATATTCCTGTGCATGAGTTTGGAGCTGCTTTAATAT 725
Db 662 CTGTGCTCTTGACAGTTTATACGACATTCAGCAGATGAAATTTGGGCGAGCACTCTGA 721
Qy 726 CTAAGAAATATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 785
Db 722 GAGGAAATGTTGATGCTGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 781
Qy 786 AGACGAGGTACGCTTAATAGAAATAGGCGCAACTTTCAAAAGAGAGAGGTGATGATGTT 845
Db 782 ATTCTGATGCTAGTCTTGACGACATAGGCGCTTTCTCTCGAGAGAGGCGATATGTTGA 841
Qy 846 CTTTCTTCTTCTGCTGATGAAAGTACTTTAAATATATAGTCAATAATACAAAATATCTG 905
Db 842 ACTTTTCTTTGTAGCAGAGATGACTTTAAATATATCTCATTTCTTATAGTAATGCTGTTA 901
Qy 906 ATTATGATGTTAAATCTTACTTCTCTGCTCTCTAGTAGAATAGTTTACTTTAAAGNAATTT 965
Db 902 AGTATGTTGTAAGACTTACTTCCCGCTTTCTAGTAGAAGATGATCAATGAGGAGTTT 961
Qy 966 TAGTCACTAGGGTTAATCTTGGTTTGTAAATTTACAAAGTAGATACCTTATATCTGT 1025
Db 962 TGGTAATCTAGGTTAAATCTTGGTTTGTAAAGTTTCAAGGTTAGATACCTTTGCTACTAT 1021
Qy 1026 ACAAGAGTTTAGACAAGTGTGATGATGATGATGATGATGATGATGATGATGATGATGAT 1085
Db 1022 ATAGAGTTGATACCAAGAGGTGTAGACAAGGAGCAATTTTACAGTGCATGGAAGATG 1081
Qy 1086 CTTTGTCTTCAAGAAACCTTGGCCATGTTTCAACACTGAAAGAGCAATCTTTAGAGACA 1145
Db 1082 CTGTGCATTAACAAAGACTTTTGGCATGATGATAGCAAGAAATCCTCTTAGAGGAT 1141
Qy 1146 CGGCTTCGGTTAACTTTTGGTTTCCCTAGATGAAAGACATGTTGATAGTACCGCTGTTG 1205
Db 1142 CATGCTGTTTAAATATGTTGGTTTCCAAAGATGAAAGATATGTTGATAGTACCTTTGTTG 1201
Qy 1206 AGGGTTCTATTACAGC-----AAAAAGATGACAAGGAGTGTAGTCAATTTAAATCGTG 1259
Db 1202 ACGTATCTTTACAGAACGAGGGGAAAGGTTTAGCAAGAAAGAGGTATGTCAGCAAGG 1261
Qy 1260 ACTTCGGTTTACAGTGTCTTAATCATATCAGAACATATCAAGCCAAAGCGTTAACTTACC 1319
Db 1262 ACTTCGGTTTACTGTGCTTAAATCATATTCGCATATACCAGTCCGAAGCGCTTACTTACG 1321
Qy 1320 AGAACGTATTATCTTTCGTGAGTCTAATAGNCCCGGTGATATCAATGGTGTACTG 1379
Db 1322 CCAATGATTAATCGTTGTTAGTTCGATAAGATCAAGAGTGAATAATCAATGGGGTGACTG 1381
Qy 1380 CTAGGTCTGAATGGATGTAGATAAGCAATTTCTTCAACCCCTTGTCAATGACTTTCTTCT 1439
Db 1382 CGCGCTCAGAGTGGATGTGATAAGGCTTTTGTGAGTCCCTGTCAATGACTTTTTTCT 1441
Qy 1440 TGCAGATCAAGCTGGCTGCGCTTCAAGACCATATAGTAAATGGGAAAGTTTCGGTGTGG 1499
Db 1442 TGCAGACCAATGGCCATGCTCAAGGATGACCTCGTGGTTTCAGAAATTCGAAGTGCAT 1501
Qy 1500 ATAGACCACTTCTGAATCTATTATGGGATCAGGTGGGCAATTTTGGAAAGCTTTCC 1559
Db 1502 CCAATCGCTCACTGATATGTTCTGGATGAGATTAAGTCTGCTGCTTTTCAATGTTTTC 1561
Qy 1560 CCACTATCAAGAGAGATTGGTGAGCAGGAAATTCCTGGATGTAAGTGAAGATGCTCTGA 1619

Db 1562 CTACAAATCAAGGAGAGGTTGATTAACAAGAAACTCATAACTGTTTCGAAAAAGGCTCTTG 1621
Qy 1620 AGATCAAGATCCCAGATCTGTATGTATCATGGAAGAGAGGTTTCGTAGCTGAATACACCA 1679
Db 1622 AAATTAAGTACCTGATTTGTATGTAACTTTCCACGATAGATTGGTTAAGAGTACAGT 1681
Qy 1680 AGTCTGAGGAGTTACCGCATCTAGATATCAAGAAAGCTTTAGAAAGCTGAGCAAAATGT 1739
Db 1682 CTTGCGTGGAAATGCCGGTACTCGACGTTTAAAAAGAGCTTTGGAAGAGCAGAAAGTATGT 1741
Qy 1740 ACACGCGGTTATCAGAAATATCTATCTCTAAAGGCTGCTGATATTTCCGATATCCGNACT 1799
Db 1742 ACAATGCTTTGTCAAGAACTCAATTTCTTAAGACAGTACAAAGTTTGATGTTGTTT 1801
Qy 1800 TCAAGACATGTGCAAGGCTTTAGATGTTTAGTCTCTGATGTGGCAGCACGAGTAATCGTTG 1859
Db 1802 TTTCCCGGATGTATAATACATTAGCGGTAGATCCATTGTTGGCAGCAAAAGGTAATGTTAG 1861
Qy 1860 CAGTGGCCGAGATAGAGCGGTTTAACTCTTACTTTTGTATAGCCAAACGAGGAGAAATG 1919
Db 1862 CTGTGGTTTCAAAATGAGAGTGGTTTACCTTAACGTTTGAAGGCTTACCGAAGCAAAATG 1921
Qy 1920 TGGCTAAGGCTCTTAAAGCAGCGGCTCTGAGCGCTGTGTATGT-----CTTGAAC 1970
Db 1922 TCGCACTTGCATTGCAACCGACAAATTAATCAATCAAGGAGGAGGTTTCGTTGAAGATTGTT 1981
Qy 1971 CGACATCCGAGAGGTGAAGTAAATAATTTTCTATTCCTGAGAAAGGAGGATGCGCTG 2030
Db 1982 CGTCAGACGTGAGTGTCTCAATCAAGAAAGCGGTTTCAAAATCAGAGATTTCTATGTC 2041
Qy 2031 TGTGTGCAAGAAAGTCAATGTTTGAAGAAATGCTTAACTTAGAGCACCGAGGAGTTGGAGTCCC 2090
Db 2042 TTGGTCTACAGGCAACACAGTGTCCGATGAGTTTCAAGAAAGTACAGAAATCGAGTCTGT 2101
Qy 2091 TCACGATTTCCATAGGCTTCGGTGGATGATGATTAACAAGCAAAATGCGATCGGTTG 2150
Db 2102 TGCAGCAGTTCCATATGTTATCCACAGAGACGATTTATCCGTAAACAGATGATGCGATGG 2161
Qy 2151 TCTACACTGCTCACTCAAGTTTCAACAAATGAAGACATGTTGAGACAGTTTGGCAGCTT 2210
Db 2162 TGTATCTGCTCGCTTAAAGTTTCAACATGCAAGACGATTTATTTAGACAGCTTGGTAGCCT 2221
Qy 2211 CGTTGTCCGCACTGTATCAAACTCTATGCAAGTCACTAAAGGATGAAGTCCGGTATGAT 2270
Db 2222 CGCTCTCTGCTCGGTTATCAAACTGAAGAGATTAATCAAGACACAGCTGCTATAGATC 2281
Qy 2271 CTGATTCAGGAGAAAGTTGGTGTGTTGGATGTCACTTTGAAAAGTGGCTCTCAAC 2330
Db 2282 TCGAGACTAAGGAAATTTTGGAGTCTACGACGTGTGCTTTAAGAAATGGTTGGTGAAC 2341
Qy 2331 CTGCGCCCAAGGTCATTTTCATGGGAGTTGCTCGATTACAAAGGGAATGTTTACTG 2390
Db 2342 CTCATCAAAAGACATGCTTGGGTTGGTGTGATGAGTCTCAGACTATAGTGTCTTGTG 2401
Qy 2391 CACTTCTATCTTATCAAGAGATAGAATGGTGACTGAGAGCGCATGAGAGGAGGTTGGCTG 2450
Db 2402 CGCTTCTCACATACGATGCGAGAAACATGTTGTGCGAGAGACATGCGGTAGTTCGAG 2461
Qy 2451 TATCATCTGATACATGTTATTTCTGATATTTGCAAGTCCCAAAATCTGAGGAAACAA 2510
Db 2462 TGAGCTCCGAATCTTTGGTGTATTCAGATATGGGAGATAGAGCTATACGCTGTGTC 2521
Qy 2511 TGAGAGCGGTGAACCCCAACCACTCTGCAAAAGATGTTACTTGTGATCGGGTGCCTG 2570
Db 2522 TTAAGACGGTGAACCCCATATAGCAGTGCAGAGTTTACACTTGTGATGCTGTTCTG 2581
Qy 2571 GTTGTGGAAGTACAAAGGAGATTTTGAAGATTTGATCTTGTATGAGATTTGATCTTGG 2630
Db 2582 GTTGGGAAGACAAAGAGATTTCTTTCTAGGGTCAACTTTTGAACGAGATCTAGTCTG 2641
Qy 2631 TTCTCTGAAAAACAGCTCTGCTATGATCAGAAAGGCTTAAATTCATCTGAGCTGATAA 2690
Db 2642 TACCAGGAAACAGGCTGCTGAATGATAGAAAGGGGCAACAGTTCTGCTGTTAATCG 2701

Db 4856 TAGTTAAGTACCTGTGTGATATAAAGGTTATTTCAACACATTTGTTTGGAGTAAATGGCGT 4915
Qy 4897 TTAGTCTCAGAGATACGTGCAAAATTTAGCGAGTTCATTTGATCTTTTCGAAAACAGGATGAGA 4956
Db 4916 TAGTAGTCAAGACGACGTGTAAGATTTCTGAGTTTCAATCAATTTGTCGCCCTGAGAAAT 4975
Qy 4957 TACTTCCCGCATTCATGACTAAGGTCAGAGAGTGTGTAGATATCGACTGTGACACAGATTA 5016
Db 4976 TCTTACCTGCTGTTATGACTTCGGTCAAGACGCTACGATTTTCGAAAGTTGACAAAGTGA 5035
Qy 5017 TGGCTGTTAAGNATCATAGTCTTTCTGATGTAGATTTACTTAAGAGTGTTAAGTTAGTTA 5076
Db 5036 TTGCAATGGAACACGATTCGTTATCCGATGTGGATTTGCTTAAAGGTTGAAGCTTTGTTA 5095
Qy 5077 AGAAGAGGTATGTGTGCTTAGCTGATTTGGTGTGTCTGGGAGTGGGAATCTCCCGGATA 5136
Db 5096 AGGATGGTTATGTGTTTAGCAGGGTTAGTTGTCTCCGGAGTGGAACTTCCCGGACA 5155
Qy 5137 ACTGCCGTGGTGTGCTAGTGTGTTGTATTTAGATAGAGATGAAAGAGGATGAAGGAAG 5196
Db 5156 ACTGCAGAGGTGGAGTAAGCGTTTGTGTTGATAGAGAAATGCAAGAGATGACGAAG 5215
Qy 5197 CAACGCTCGGTGCTATCACGCCCTGCTTGCAGAAAGAAATTTTCTTTTAAAGCTAATCC 5256
Db 5216 CAACACTTGGATCTTATAGAACCAAGTGCAGCTAAGAAACGATTTGCGCTTCAAAATGATCC 5275
Qy 5257 CTAATTTATTAACATCCGAGGATGCTGAGAGACCCCGTGGCAAGTGTAGTGAATA 5316
Db 5276 CGAATATAGCATTTACTACCCCGGATGCTGAGAGAAAGTTTGGCAAGTTTGTAGTAAATA 5335
Qy 5317 TCAAGGAGTGGGTATGGAAGAGATCTGCTCTTTATCTTTGGAGTGTGTTTCAATTT 5376
Db 5336 TTAGAGTGTGTGCCATGGAAGAGGTTTCTGCTCTTTATCTTTGGAGTGTGTTCTCAGTTT 5395
Qy 5377 GTGTAGTACATAAAATATGTAAGAAAGTTTGGAGGACGTATTTTGGAGTGTGACAG 5436
Db 5396 GTATTGTACAAATCCATATAAATTTAGGCTTTGAGAGAGAAATTAAGTGTGTCAG 5455
Qy 5437 ACAGCTGCCAATTAACACTCACTGAAAGGTTTGTGAGGAGTTCGTGGATGAAGTACCAA 5496
Db 5456 AAGGAGGACCGTTGAACTTACAGAGACGCTGTTGATGAGTTCATCGATCAGTTCCAA 5515
Qy 5497 TGGCTGTGAACACTCGAAAGGTTCCGGAACAAACAAAGAAATGTTAGGTAATAATGTTA 5556
Db 5516 TGGCTGACAGATTACGTAATTTTCGCAATCAATCTAAGAAAGGAAGTAATAAGTATGATAG 5575
Qy 5557 ATAATAAGAAATAAATAACAGTGTGAAGAGGTTTAAATTTGAGGAAATTTGAGGATA 5616
Db 5576 GTAAGAGNATGATAAAGGTTTGAATAGGAGGGAAGCTGTTTGTGAAGGTTAGAA 5635
Qy 5617 ATGTAAGTATGACGAGTCTATCGCGTCATCGAGTACGTTTAAATCAATATGCTTTATAC 5676
Db 5636 TTGGGCAGAACTCGGAGTTCATCGAGCCGAGTCTTCTCGTTTAACTATGCTTACAC 5695
Qy 5677 AATCAATCTCCGAGCAATTTGTTTACTTATCTTCGCTTACGAGATCCCTGCGAGCT 5736
Db 5696 AGTTCCAGTGCCCAATCAATAGTGTATTTAGGTTCTGTATGGCTGATCCATTAGAGTT 5755
Qy 5737 GATCAATCTGTGACAAATGCAATGGGTAAACAGTTTCAACGCAACAGCTAGGACAAC 5796
Db 5756 ACAAATCTGTACTTCGGGTTAGGCATCAGTTTCAACACACAGCTAGAACTAC 5815
Qy 5797 AGTCCACAGCAATTTCCGAGTCTGGAACCTGTGCTCTAGTATGACAGTGAATTTCC 5856
Db 5816 GGTTCACACAGCAGTTCTCTGATGTGTGGAAGACTTATCCGACCGCTACAGTTAGATTTC 5875
Qy 5857 TGCATCGGATTTCTATGTGTATAGATATAATTCGACGCTGTATCGCTTGTATCACCGGTT 5916
Db 5876 TGCTATGTTTCAAGTTTTCGATATAATGCGGTGTAGATTTCTCTAGTGTGCGCACT 5935
Qy 5917 ATTAATAAGCTTTGATACTAGAAATAGAAATAATAGAGGTTGATAATCAACCCGACCGAA 5976

Db 5936 TCTCGAGCCTTTGATACTAGGAACAGGATAATAGAGTTGAAAATCCGCAAAATCCTAC 5995
Qy 5977 TACTACTGAATCGTTAAACCGACTCAGAGGGTAGAGTACTGTACTGTAGCTATAAGGGC 6036
Db 5996 AACTGCCGAGACGCTTGATCGACGAGCGGGTAGATGATCGACGCTGCGCATTAGGGC 6055
Qy 6037 TTCAATCAATAAATTTGGCTAATGAATCGTTTCGTGGAACCTGGCATGTTTCAATCAAGCAGG 6096
Db 6056 CAGTATAGTAACTCATGATAGTATGTTCTGGCACGGGAATGTACAAATCAAGCTCT 6115
Qy 6097 CTTTGAGACTGCTAGTGGACTTCTCTGGACACAACTCCGGCTACTTAGCTATTGTTGTG 6156
Db 6116 GTTCGAGAGCGGAGTGGACTCACCTGGGCTACAACTCCTTAAAC----- 6160
Qy 6157 AGATTTCTTAAATAAAGTCGCTGAAGACTTAAATTTAGGGTGGCTGATACCAAAATCA 6216
Db 6161 ATGATGGCATAAATAAGTTGGACGAACATTAACAGTCC---GTGGCAGTAGTACGATAACTC 6217
Qy 6217 GCAGTGGTTGCTCGTCCACTTAATAATAAGATTTGTCATATCTGGATCGAACAGATTTAAAC 6276
Db 6218 GTAGTGTTTTCCCTCCACTTAATCGAGGGTGTGCTGGGATGGAACGCAATTAAT 6277
Qy 6277 CATGTGATGTGTATCTGTGGTATCGGCTAA--AACATCGAGAGGTTTGAATCTCTCCC 6335
Db 6278 ACATGTGTGACGTGTATTTGCGAAACGACGTAATTTATTTTTCAGGGGTTTGAATCCCCC 6337
Qy 6336 TAACCGCGGTAGCGGCCCA 6355
Db 6338 GAACCGCGGTAGCGGCCCA 6357

RESULT 15

PPPSLF LOCUS DEFINITION ACCESSION VERSION KEYWORDS SOURCE ORGANISM REFERENCE AUTHORS TITLE JOURNAL MEDLINE PUBMED COMMENT FEATURES source

PPPSLF 6357 bp ss-RNA linear VRL 13-JUN-1996
Nucleotide sequence of the genomic RNA of the tobamovirus Pepper
mild mottlevirus, strain S (PMV-S).
M81413.1 GI:333293
coat protein; replication-associated protein.
Pepper mild mottle virus
Pepper mild mottle virus
Viruses; ssRNA positive-strand viruses, no DNA stage; Tobamovirus.
1 (bases 1 to 6357)
Alonso, E., Garcia-Luque, I., de la Cruz, A., Wicke, B.,
Avila-Rincon, M.J., Serra, M.T., Castresana, C. and Diaz-Ruiz, J.R.
Nucleotide sequence of the genomic RNA of pepper mild mottle virus,
a resistance-breaking tobamovirus in pepper
J. Gen. Virol. 72 (Pt 12), 2875-2884 (1991)
92113528
1765765
Original source text: Pepper mild mottle virus (strain S) cDNA to
genomic RNA.
Location/Qualifiers
1..6357
/organism="Pepper mild mottle virus"
/mol_type="genomic RNA"
/strain="S"
/db_xref="taxon:12239"
70..4908
/note="183K ORF; internal stop codon"
/codon_start=1
/transl_except="(pos:3421..3423,aa:OTHER)"
/product="replication-associated protein"
/protein_id="AAB02334.1"
/db_xref="GI:333295"
/translation="MAYTQATNAALASTLRGNPLVNDLANRLYESAVSQNHDR
RPKNFLRSISEQTLLKATPEFQITFNQNAVSLAGLSLELEYLMMQIPY
STTYDIGNFAHRPFKRDYVHCMPNMDLRVHRNAQKDSIELYSLKAKKKVIP
PYQKPFKDYDDPSQVVKSPFQHCQVSGHCTDKYVALHSLYDIPADFGAALLR
RNHVCAAFHSENLILSDYVSLDDIGAFREGLMNFSAESTLNTHTSVSV
LKVCYKTFPASREVMKFLVTRVNTWFKFSRLDTFVLYRGVTHRGVDKEFYSA
MEDAHYKKTLLAMNSRILLEDSSSVNYWPKMKDMVIVLPFVLSQNEKRLARKE
VMVSKDFYTVLNIHRTQSKALTYANLVSFVESIRSKVITNGVTARSEWDVDDKALLQ


```

LSWTFPLQTKLAWLKDDLVQKFOVHKSLSLTRYVMDLTAAPHNCPTTKERLNKK
LITVSEKALEIKVPDLYTYFHDRLVKEYSVSEMPVLVDKSLSEAEVWMLNALSEISI
LKDSKFDVDFVFRMCDNTLGVDPVLAQVWVAVVSNESGLTLTFEPTBANVALALQP
LITVSEKSLIKHVVSSDSESSIKVYKVEISMLGTNTVDSFORSFTEISLQOPH
MVSTETIRKQKHAMVYTGPKVKQCKNYLDSLVASLSAAVSNLKKIIKDTAIDLDET
KEKFGVDVCLCKWLVKPLSKGHAWGVWDSYKFCFVALLTYDGENIVCGETWRRVAV
SSSLVYSDMGKIRATRSVLKDGEPHISAKVTLVDGPGCGKTKELSRVNFDEDLV
LVPKQAAEMIRRANSSGLIVATKENVRTVDSFLMNYGRGPCQYKRLFDSEGLMLHP
GCNFFLVGMSLSEAFYIGDTQIQIPIYNIRVATFPYKHLSQLSEVDAVETRRITLRCFA
DIITFFLNQKYGEGVCMCTSVRSVSHVEIQAAMVNPVSKPLGKVIITFTQSDKSLLL
SRGEYDVHTVYHQGTFFEDVSLVRLTTPVGIISQSPHLLVSLSRHRSYKYITVW
LDAVSVLRDLCECVSSYLLDMYKVDVSTQXQLOQIESVYKGVNLFVAAPKTDGVSDMOY
YDKCLPGNSTILNEYDAVMOJRENSLVNDCVLDMSKSVLPRESESTTLKPVIRTA
AEKPRGGLLENLAVMKNFNAPSPELGVVDIEDTASLVVDKPEFDAYLKEKKKPKNI
PLLSRASLERWIKQSKSTIGQADDFDIDLPVADQTRRMKIQQKQRLDLSIQTSEIP
ALQTVIYHKSINKALPGFVSELTROYELTIDSSRFMYTRKPTQIQIEBFSDLSNSV
PMDILBDSIKYDKSQNEHCAVEYIEIKRWGLDDFLAEVWKHGRKTLTKDYTAGIK
TCLWYQRKSGDVTFITGNTIIIAACSLMPLMERLIKGAFCGDDSLPYPKGTDPDI
OQGANILMNFPEAKLPRKRYGYCFGRVYIHHDRGCIYVYDPLKLSKLGAKHINKREHL
EEFRTSLCDVAGSLNANCAYTHLNDVAGEVITAPLGSFYRALVYKLCRKLPTLFL
LE"
70..3423
CDS
/notes="126K ORF"
/codon_start=1
/product="replication-associated protein"
/protein_id="AAB02335.1"
/db_xref="GI:333294"
/translations="MAYTQATNAALASTLRGNPNLVNDLANRRLYESAVQCNADR
RPNVFLRGISSEQTIIATKAYPEFQITFYNTQNAVHSLAGLSRLSELSYLMQIPIYG
STYDIGNPFAAMQGRVDVHCCPMNDLRDVMRHNAOKDSIELVLSKLAOKKVIPI
PVQKPCFDKYDQPSVCSKPKQHCQEGYSHCDTKYVAVALHSLVDIPADERGAALLR
RNHVCIYAAHPSENLLSDSYSLDDIGAPFSREGDMLNPSFVAESTLNTYHSYNSV
LKVKCTYTPASSREMYKFLVRVNTWPKFSLDITFVLRGVYHVRGVDKEQFYSA
MEDAHYKXTLANMNERILLEDSSVNTYFPRKMDMVIPLFDVSLQNEGRFLARKE
VMSKDPVYTVLNIHRTYQSKALTYANVLSFVRESIRSVINGVITARSBDVDKALLO
LITVSEKALEIKVPDLYTYFHDRLVKEYSVSEMPVLVDKSLSEAEVWMLNALSEISI
LKDSKFDVDFVFRMCDNTLGVDPVLAQVWVAVVSNESGLTLTFEPTBANVALALQP
LITVSEKSLIKHVVSSDSESSIKVYKVEISMLGTNTVDSFORSFTEISLQOPH
MVSTETIRKQKHAMVYTGPKVKQCKNYLDSLVASLSAAVSNLKKIIKDTAIDLDET
KEKFGVDVCLCKWLVKPLSKGHAWGVWDSYKFCFVALLTYDGENIVCGETWRRVAV
SSSLVYSDMGKIRATRSVLKDGEPHISAKVTLVDGPGCGKTKELSRVNFDEDLV
LVPKQAAEMIRRANSSGLIVATKENVRTVDSFLMNYGRGPCQYKRLFDSEGLMLHP
GCNFFLVGMSLSEAFYIGDTQIQIPIYNIRVATFPYKHLSQLSEVDAVETRRITLRCFA
DIITFFLNQKYGEGVCMCTSVRSVSHVEIQAAMVNPVSKPLGKVIITFTQSDKSLLL
SRGEYDVHTVYHQGTFFEDVSLVRLTTPVGIISQSPHLLVSLSRHRSYKYITVW
LDAVSVLRDLCECVSSYLLDMYKVDVSTQ"
4909..5682
/misc_feature="cell to cell spreading of the virus"
/notes="30K ORF"
/codon_start=1
/protein_id="AAB02336.1"
/db_xref="GI:333296"
/translations="MALVKDDVKIISFINLSAEKFLPAVMTSVKTVIRISKVDKVI
MENDSLSDVNLGKVLKVDGYVCLAGLVSGEMNLPDNCRGVSVCLVDKRWQRDE
ATUGSTRISAARKFAPKLPNTSITADAKRWQVQVLRNIRGVAMEKGFCLPSLEFV
SVCIVHKSINKLGRBKITSVSEGPVELTEAVVDFEISVPMADRLRFRNQSKKGS
NKTVGRNDNKGLEKGLFDKVRIGQNSSESDAESSP"
5458..5517
/misc_feature
/notes="assembly origin"
5685..6158
/misc_feature
/notes="3'-end noncoding region (tRNA-like sequence)"

```

```

ORIGIN
Query Match      39.8%; Score 2528.4; DB 14; Length 6357;
Best Local Similarity 63.6%; Pred. No. 0;
Matches 4056; Conservative 0; Mismatches 2266; Indels 58; Gaps 11;

QY      6  TTTAATAGTTTTTCGACAAACAACAATTTAAACAAAAACAACATATTACAACAACAACA 65
DB      6  TTTTTCACAAATTTAAACAAACAACAACAACAACAACAACAACAACAACAATACAA 64

QY     66  CAACTAATGGCACACATACATCTATTAATTTAGCAACGCCCTTCTTTGAAAGCGTGAGTGA 125
DB     66  CTACAAATGCTTACACACACAACAAGCTTACCAACGCCGCATTTAGCAAGTACTCTCCGAGGA 124

QY    126  AAAAACTCTCGTTAATGACCTTTCAGAAAGGCGCATGTACGATCGGCGCGTGAAGAAT 185
DB    126  AAAAACTCTCGTTAATGACCTTTCAGAAAGGCGCATGTACGATCGGCGCGTGAAGAAT 185

QY    125  ATAACCCCTTGGTGAACGATCTTGTGTAATCGGAGACTGTACGAATCAGCGGTGCGAACAT 184
DB    125  ATAACCCCTTGGTGAACGATCTTGTGTAATCGGAGACTGTACGAATCAGCGGTGCGAACAT 184

QY    186  TTAAGCCCGCGACCGTAGACCAACAAAGGTCAACTTTTCCAAAACTATTAGCGAAGAGCAAA 245
DB    186  GCATGCAATGACCCGACGCGCCCAAGGTTAATTTTTTAAGTTCGATAGCGAAGAGCAGA 244

QY    246  CGCTTCTAGTCTCCAAACGCGTACCCGAGGTTCAGATTAACCTTTTATAATCTCAAAATG 305
DB    246  CGCTTATCGCAACTAAGGCTTACCCCTGAGTTCCAAATCAAGTTCTACAACACGACGACG 304

QY    306  CCGTACACAGTTTGGCTGGAGGTTTGAGAGCATTTAGAAATTTGGAAATCTGTATCTCAAG 365
DB    306  CCGTACACAGTTTGGCTGGAGGTTTGAGAGCATTTAGAAATTTGGAAATCTGTATCTCAAG 365

QY    366  TTCCTCTATGATCGCGACATATGATATAGTGGGAACCTTTCAGCACACATTTCTTCAAG 425
DB    366  TTCCTCTATGATCGCGACATATGATATAGTGGGAACCTTTCAGCACACATTTCTTCAAG 425

QY    426  GCAGGGAATACGTGCAATTCGTGTATGCCCAATCTCGACATACAGAGATATAATGAGGACG 485
DB    426  GTCGTGACTACGTTCAATTCGTGCAATGCTTAAATGCACTTACGTGAGTCACTCGTCACA 484

QY    486  AAGGACAAAGGACTCAATTTGAGATGATATTTGTCCAGATTTGTCTCTGTTCTTAAACAGTAA 545
DB    486  ATGCTCAAAAGGATAGTCAATTTGAGATGATATTTGTCCAGATTTGTCTCTGTTCTTAAACAGTAA 544

QY    546  TTCCTGAGTTTCAAGGGAGGCTTTTAAACAGGTATGTCAGAGCTCCAGAGCTCCACAGAGTCTGCT 605
DB    546  TACCGGCATATCAAAAGGCCATGCTTTGATAAATACACGACGATCCGCAATCACTAGTAGTGT 604

QY    606  GCTCTAAAACTTTTCAGGAATTTGCGAATACATCCCGCAGAGATAGTGGTAGAAGATACG 665
DB    606  GCTCGAAACCTTTTCAGCACTCGGAAGGCTTTGCG---ACTGCACGATAAAGTATACG 661

QY    666  CTGTTGCTCTGCACAGTTTGTATGATATTTCTGTGTGATGAGTTTGGAGCTGCGTTAATAT 725
DB    666  CTGTCGCTTTGCACAGTTTATACGACATTTCCAGCAGATGAAATTTGGGGCAGCACTTCTGA 721

QY    726  CTAAGAAATACATGATGATGTTATGAGCTTCCATTTTGGCAGAGCATTTATTACTAGACC 785
DB    726  GGAGAAATGTTTCATGCTCTGATGTCCTTCCACTTTTCTGAGAAATCTTCTTTTGAAG 781

QY    786  AGACGGAGGTACGCTTAATGAAATAGGCGCAACTTTTCAAAAAGAGAGGTGATGATGTTT 845
DB    786  ATTCTGATGTCAGTCTTGACGACATAGGCGCTTTCTTCTCGAGAGAGGGCGGATATGTTGA 841

QY    846  CTTTTTCTTTGCTGATGAAAGTACTTTTAAATATATAGTCATTAATATCAAAAATATCTTGC 905
DB    846  ACTTTTCTTTTGTAGCAGAGTACTTTTAAATATATATCTCATCTCTATAGTAAATGTGCTTA 901

QY    906  ATTATGATGTAATCTTACTTTCTGCTTCTAGTAGAATAGTCTTATTAAAGGATTTT 965
DB    906  AGTATGTTGTGAAGTACTTCTCCCGCTTCTAGTAGAAGTGTACATGAAGGAGTTT 961

QY    966  TAGTCTACTAGGGTTAATCTTGGTTTGTGTAATTTTACCAAAAGTAGATACCTATATCTGT 1025
DB    966  TGGTAACTAGGGTAAATCTTGGTTTGTGTAATTTTCAAGGTTAGATACCTTTGTACTAT 1021

```

1026 ACAAGAGTGTAGACAAGTAGGGTGTGATAGTACAGTTCATGAGCGGATGGAAGACG 1085
1022 ATAGAGGTGTATACCAACAGAGGTGTAGACAAGGAGCAATTTTACAGTCAATGGAAGATG 1081
1086 CTTTGGCTTACAAAGAAACCTTGGCCATGTTCAACACCTGAAAGAGCAATCTTTAGAGACA 1145
1082 CTTGGCATTTACAAAAGACTTTGGCGATGATGAATPAGCGAAGAATCCCTCTAGAGGATT 1141
1146 CGGCTTCGGTTAACTTTTGGTTCCCTAAGATGAAGGACATGGTGATAGTACCGCTGTTTG 1205
1142 CATCGTCTGTTAAATTATTTGGTTTCAAGATGAAGATATGGTGATAGTACCTTTGTTG 1201
1206 AGGGTTCTATATACAGC-----AAAAAGATGACAAGGAGTGAGGTCAATTTGTAATCGTG 1259
1202 ACGTATCTTTACAGAACGAGGGGAAAGGTTAGCAAGAAAGAGGTCAATGTCAGCAAGG 1261
1260 ACTTCGTTTACACAGTGTCTTAATCATATCAGACATATCAAGCCCAAGCGTTAACTTACC 1319
1262 ACTTCGTTTACTGTCTTAATCATATTCGCACATACCACTCGAAAGCGCTTACTTACG 1321
1320 AGAAGTATTATCTTTCCGTGGAGTCTATAAGATCCCGGTGATTAATCAATGGTGTACTG 1379
1322 CCAATGTATATCGTTCTGCTGATGATAGTCAAGTGTATCAATGGGTGACTG 1381
1380 CTAGGTCTGAATGGATGTATGAATAAGCAATTTCTTCAACCTTGTCAATGACTTTCTTCT 1439
1382 CGCGCTCAGAGTGGGATGTGATAAGGCTTTGTTGCAGTCCCTGTCAATGACTTTTTTCT 1441
1440 TGCAGCTAAGCTGCTCGCTTCAAGACGATATAGTAATGGAAGGTTTCGGTGTCTGG 1499
1442 TGCAGCCAAATTTGGCCATGCTCAAGGATGACCTCGTGTTCAGAAATTCAGATGCAATT 1501
1500 ATAAGACCACTTCTGAATTTTGGATGAGGTGGGCAAAATTTTGGAAACGTTTTC 1559
1502 CCAATCGCTCACTGAATATCTCTGGGATGAGATTACTGCTGCTTTTCAATTTGTTTC 1561
1560 CCACTATCAAGAGAGATTGTGAGCAGGAAATTTCTGGATGTAAGTGAGAACTCTCTGA 1619
1562 CTACAATCAAGAGAGGTTGTAATCAAGAAATCTATAAATCTGTTTCGGAAGGCTCTTG 1621
1620 AGATCAAGATCCAGATCTGTATGTCTCATGGAAGACAGTTCGTAGCTGMAATACACCA 1679
1622 AAATTAAGTACCTGATTTGTATGTAATTTCCAGATAGATTGTTGAAGAGTACAAAT 1681
1680 AGTCTGAGGATTAACCGCATCTAGATCAAGAAAGACTTAGAAGAGCTGAGCAATGT 1739
1682 CTTCCGTTGGAATGCCGTAAGTCTGACGCTTAAAGAGCTTGAAGAGAGCAAGTGAATG 1741
1740 ACGAGCGTTATCAGAAATTAATCTATCTTAAGGTGCTGATTAATTCGATATCGCAAGT 1799
1742 ACAATGCTTTGTCAAAATCTCAATCTTAAAGACAGTGAACAAGTTTGTATGTTGATGTTT 1801
1800 TCAAAGACATGTGCAAGGCTTTAGATGTAGTCTGATGTGGCAGCACAGTAAATCGTTG 1859
1802 TTTCCCGGATGTGTAAATACATTAGGCGTAGATCAATTTGGTGGCAGCAAGGTAATGTAG 1861
1860 CAGTGGCCGAGATPAGAAAGCGTTTAACTTTACTTTTGATAAGCCAAACCGAGGAGAAATG 1919
1862 CTGTGGTTTCAAAATCAGAGTGTGTTGACCTTAAACGTTGAGAGGCTTACCAGAAAGATG 1921
1920 TGGCTAAGGCTCTTAAAGACGCGGCTGAGGCGGTGTTGTTCTT-----CTTGAAC 1970
1922 TCGCATTTGCTGCAACCGCAATTAATCAATCAAGGAGGAGGTTGTTGAAGATTGTGT 1981
1971 CGACATCCGAGAGGTGAACGTAAATAATTTTCTATTGCTGAGAAAGGAGATTCGCTG 2030
1982 CGTCAGACGTAGGTGAGTCTCTCAATCAAGAGAGTGTTCGAATATCAGAGATTTCTATGC 2041
2031 TGTGTGAGAAAGTCTAGTTTTCAGATCTTAATCTTAGAGCACCGAGGAGTTGAGTCCC 2090
2042 TTGGTCTAAAGGCAACACAGTGTCCGATGAGTTTCCAAAGAGTACAGAAATCGAGTCTG 2101

2091 TCACAGATTTCCATAAGGCTTGGTGGATAGTGTGATTACAAAGCAAAATGGCATCGTTG 2150
2102 TGCAGCAGTTCCATATGGTATCCACAGAGCAGTATATCCGTAAACAGATGCTGCCATGG 2161
2151 TCTACACTGGCTCACTCAAAGTTCAAAATGAAGAACTATGTGGAAGTTTGGCAGCTT 2210
2162 TGTATACTGTGTCGGCTAAAAGTTCAACAAATGCAAGAACTATTTAGACAGCGCTGGTAGCCT 2221
2211 CGTTGTCCGCCACATGTATCAAAATCTATGCAAGTCACTAAAGGATGAAGTCGGGTATGATT 2270
2222 CGTCTCTGTGTCGGTATCAAACTGAAGAAGATAATCAAAAGACACAGCTGCTATAGATC 2281
2271 CTGATTTCCAGGAGAAAGTTGGTGGTGGTGTCTCTTTGAAAAAGTGGCTCTCAAAAC 2330
2282 TCGAGACTAAGGAAATTTGGAGTCTACGACGTGTGCCCTTAAGAATGGTTGGTGAAC 2341
2331 CTGCGGCCAAAGTCAATTCATATGGGAGTTGTCTCTGATTAACAAGGGGAAAAATGTTTACTG 2390
2342 CTCTATCAAAAGGACATGCTTGGGTGTGTGATGACTCAGACTATAAGTGTCTTGTGTTG 2401
2391 CACTTCTATCTTATGAAGGAGATAGAAATGGTGAAGTCTGAGAGCGACTGGAGAGGGTGGCTG 2450
2402 CGCTTCTACATACGATGGCGAGAACATTTGTGTGGGAGAGACATGGCGTAGTGCAG 2461
2451 TATCATCTGATACAATGGTATATTTCTGATATTGCAAAAGCTCCAAAATCTGAGGAAAAACAA 2510
2462 TGAGCTCCGAATCTTTGGTGTATTCAGATATGGGAGAGATAGAGCTATACGCTCTGTGC 2521
2511 TGGAGACCGGTGAACCCACGAACTCTCTGCAAGATGTGCTTGTGGATGGGTGGCTG 2570
2522 TTAAGACCGGTGAACCCCATATAAGCAGTGCAGAGGTTACACTTGTGTGATGGTGTCTG 2581
2571 GTTGTGGAAGTACAAAGGAGATTTTGAAGATTTGATCTTCATGAGGATTTGATCTTGG 2630
2582 GTTTCGGAAGACAAAGGAGATCTTTTCAGGGTCAACTTTGACGAGATCTAGTCTTGG 2641
2631 TTCCTGGAACAAAGCTGTCTGTATGATCAGAAAGGGCTAAATTCATCTGAGCTGATAA 2690
2642 TACCAGGAAACAGGCTGCTGAAATGATAAGAAAGGSCAAACAGTTCTGTTTAAATCG 2701
2691 GAGCCACATGACATGTGAGAACGGTAGATTCTACTTCTATGCA-----TCCAAAC 2744
2702 TGGGACCAAGGAGATGTAGGCGGTAGACTCTTCTTATGAAATTAACGCTCAGGTC 2761
2745 CGCGATCAACAAGAGGCTTTTATTTGATGAAGGTTGATGCTGCACACCGGTCTGTGTTA 2804
2762 CGTGCACATACAAAGGCTGTTCTGGATGAAGTCTAATGTTACACCTGTGTTGTTA 2821
2805 ACTTCTGTGTCTTATCTCTGTGTCGACATCGCATACATTTACGAGATACACAGCAGA 2864
2822 ATTTTCTGTGTCATGCTCTATGCTCCGAGGCTTTTGTGTTATGAGACACCCAGCAGA 2881
2865 TTCTCTTCAATTAACAGATTTCAGAAATTTCCGTATCCCAACCATTTTGGAGCTGCAAG 2924
2882 TTCTTACATCAACAGATTGCACTTTTCCCTATCTTAAAGCATTTGAGTCAACTCGAGG 2941
2925 TGGATGAAGTTTCAGATGAGGAGGACCACTGAGATGCCAGGTGATGTGAAATTTTTC 2984
2942 TCGATGCTGTTGAACCTCGCAGAACCAAGTTTCGGGTGTCAGCTGATATCACTTCTTCT 3001
2985 TACAATCGAAGTACGAAGGCGGTGACAAACATTCATCTGTGTACCAAGTCGGTCTCAT 3044
3002 TGAATCAGAAGTACGAAGGCGCAAGTTATGTGCACATCAAGTGTTCACGCTCGGTGTCA 3061
3045 CTGAGATGATAGGCGGTAAAGGAGTACTAAACAGTGTTCCAAACCATTAAGGGAAGAA 3104
3062 ACGAGTCTCAAGGTCCAGCGGTATGAATTCAGTGTCTTAACCATTTAAGGGAAGG 3121
3105 TTGTAACCTTCACTCAGGCTGATAAATTTGAGTTAGAGGAGAGGCTTATAGAAATGTGA 3164
3122 TGAATACATTCAGTCAAGTCAAGTCAATTTGCTGCTCTCGAGGGGTACCAAGATGTGC 3181
3165 ACACCGTTTCATGAGATCCAAGGAGAAACCTTTGAAGATGTGTCTGCTGCTCAGATTGACGG 3224

|||||
3182 ATACCGTTCAATGAGTGCAGGGGAAACGTTTGAAGACGCTCTCACTAGTAGGCTGACGC 3241
QY
3225 CAACTCCAAGTCTGATTTCAAGTCTTCCCGCATGTTCTAGTGGCTCTGACTAGAC 3284
Db
3242 CAACACCCGTGGGAATAATTTCAAGCAGAGTCCGACACCTGTTGGTCTCATTTGCTTAGGC 3301
QY
3285 ACACAAGAGCTTCAATATTAACACCGTAGTGTAGATCTTTAGTACAGATAATTTAGTG 3344
Db
3302 ATACAAGGTGCGATCAAAATATTACACAGTTGTGTAGATGAGTCGTTTCAAGTCTTAGAG 3361
QY
3345 ATTGTGCTTTCTTAAGCTCTTCTTTTAGAATATATATGTTAGAGCAGGTAGTAGAT 3404
Db
3362 ATCTGAGTGTGAGTAGTTACCTGTTAGATATGTACAAGTTGATGTGCGACTCAAT 3421
QY
3405 AGCAATTAAGATGAGTGTGTTCAAGGTCAATATCTCTTTGTTGGCAACACCTTAAT 3464
Db
3422 AGCAATTAAGATGAGTGTGTTCAAGGTCAATATCTCTTTGTTGGCAACACCTTAAT 3481
QY
3465 CAGGAGACTTCCAGATCTACAGTCTTATTAAGATGTATGCTCCCTGCTGATAGTACTA 3524
Db
3482 CAGGAGATGTTCTGACATGCAATATTAACGACAAGTGTGTCGCGGAAACAGTACTA 3541
QY
3525 TACTTAAAGATGATGCTGTTTACCATGAGGTTACGTGATATAGTCTTAATGTGAAGG 3584
Db
3542 TACTCAATGATGATGCTGTTTAACTATGCAATACGAGGATAGTGTGATGTGCAAGG 3601
QY
3585 ATTGTGTTCTTGTATTTTCCAAAGTATTCGATGCAAGGAGGTGAACCAATGTCATG 3644
Db
3602 ATTGTGTTGATATGTCGAAATCGTGCTCTCCGAGAGAACTCGAGACGACATTTGA 3661
QY
3645 AGCCAGTTTTGGTACCGCGGCGGACCGCAAGGCTCGAGGACTCTCGAAATCTGG 3704
Db
3662 AACCTGTGATAGGAGTCTGCTGTAAGAAACCTCGAAACCTCGAATTTGTTGGAAATTTGG 3721
QY
3705 TTGCAATGATTAAGAAATTTCAACGCAACGACCTGACGGGAGGATTTGACATTTGAGA 3764
Db
3722 TCGGATGATCAAGAAATTTCACTCTCCGAAATTTAGTGGGTTGTTGATCGAAG 3781
QY
3765 GCACCGCTCTCTGTAGTAGTAACTTTTGTGATGACTATTTTATTAAGAAAGAAAT 3824
Db
3782 ACACCGCTCTCTGTAGTAGTAACTTTTGTGATGACTATTTTATTAAGAAAGAAAT 3841
QY
3825 ACACAAATATTTGCTGAGTGTACGAGGATTCAGAGATTCAGATGAGATGGTTGGAACA 3884
Db
3842 AACCAAAATATACCT---CTGCTTTCAAGGCGAGTTTGGAAAGATGGATCGAAAGC 3898
QY
3885 GGAAGAGTACTATTTGGACGACTTGGCTAACTACAAATTTACAGATCTGCGGCCATCG 3944
Db
3899 AAGAGAGTCAACATTTGGCCAGTTGGCTGATTTTGAATTTTACAGCCGTTG 3958
QY
3945 ATCAGTACAGCAATGATCAAGGCTCAACCAAAACAGAAATTTGGACCTTTCAATTCAGA 4004
Db
3959 ATCAATACAGGCACATGATCAAGCAGCAGCGAAACAGCGTTTGGATCTTAGTATTTCAA 4018
QY
4005 ATGAATACCTGCTCTGCAAAACATTTGTCTACATTCGAGCAGATCAACGTTATTTTGG 4064
Db
4019 CTGAATACCGGCTTTGCAAACTATTGTGTATCATAGCAAGAAATCAATTCGCTTTTGG 4078
QY
4065 ---CCGCTTTCTCAGAGCTTACAAAGTGTGCTGCTCGAGGCAATTTGATTTCTAAGAGTTTC 4121
Db
4079 GTCTCTATTTTCAAGATTTAACAAGACAGTCTGTAGAGCAATTTGACAGTTCAAGATTTCA 4138
QY
4122 TTTTCTTTTACTAGGAAATCTCAGAAACAGATTTCAAGAAATTTTCTCGGATCTGACTCGC 4181
Db
4139 TGTTTTATACAGGAAACCGCTTACACAGATCGAAGAAATTTTCTCAGATCTGACTCTA 4198
QY
4182 ACCTTCTATGAGTGTGTAGAACTGGATTTTCTAAGTATGATATAGTCAAGAACGAT 4241
Db
4199 ATGTTCTATGACATATTAGAGCTAGACATTTTCCAAGTATGACAAATCACAGAACGAAT 4258
QY
4242 TTCAATTTGCTGTAGTAGTGAATATGGAAGATTTGGTCTCAATAGTATTGTCGCG 4301
|||||

Db
4259 TTCAATTTGCTGCTGAGTATGAGATTTGGAAAGGTTAGGCTTTAGACGATTTCTTTGGCTG 4318
QY
4302 AAGTGTGGAACAACAGGCGACAGGAAACAACTTTGAAAGATTCATTTGCTGGAATCAAGA 4361
Db
4319 AAGTTTGGAAACACAGGCGATCGGAAGACAACTTTGAAAGACTACACAGCGGAATAAAA 4378
QY
4362 CATGCTGTGGTATCAAAAGGAAACGGGTGATGTGACTACTTTTCATCGGCAATACCTGTTA 4421
Db
4379 CGTGTGTTGGTACACAGGAAACGGGTGATGTGCTCACTATTTCTTTGGAACACGATCA 4438
QY
4422 TAATAGCAGCTTGTGTTGGGTTCAATTTACCGATGGAAGGTCAATAAGGTTGCTTTTT 4481
Db
4439 TTATTTGCTGATGTCTGCTCTCTATGCTACCGATGGAGAGATTGATTAAAGGTGCTTTTT 4498
QY
4482 GTGGAGACGATTTCCGTTTGTATTTTCAAGGTTTGAATTTCCCTGACATTTCACTCAT 4541
Db
4499 GTGGTGTATGATATATTTATATCTTTCCAAAGGCGACTGATTTTCCCGGATATTTCAACAGG 4558
QY
4542 GTGCTAACTCTATGTGGAAATTTTGAGGCCAAACTGTATAGAAGAGGTACGGTTACTTTT 4601
Db
4559 GCGCAAACTTCTCTGGAATTTTGAAGCCAGTTGTTTAGGAAGAGATATGTTTACTTTTT 4618
QY
4602 GTGGTAGATACATCATACCAATGATTAAGGAGCAATAGTGTATTTATGATCTCTTTGAAGT 4661
Db
4619 GCGGTAGGTACATAATTTCAACATGACAGAGGCTGTATTTGATATTTATGACCTCTAAAT 4678
QY
4662 TGATCTCCAACTTTGGGCGCAAAACATATCAAGATTTATGATCACTTAGAAGAGTTAAGG 4721
Db
4679 TGATCTCGAACTCGGTGCAAAACACATCAAGATAGAGAACTTTTAGGAAATTTAGGA 4738
QY
4722 TGCTTTTGTGCGATGTTGCTGCTCGCTCGGAACTGGTGTAGGCTTTTCGACGCTGA 4781
Db
4739 CTTCTTTTGTGATGTTGCTGGGTCGT---GAACAATTTGCGGTACTATACACATTTGA 4795
QY
4782 ACCGAGCTATCAAGAGAGGTTCAATAAACCGCGCAATTTGATGTTGCTTTGCTTTTAAATG 4841
Db
4796 ACAGCGCTGTCGTTGAGGTTATTAAGACCGCACCTCTGTTGCTTTGTTTATAGAGCAT 4855
QY
4842 TTAAACAAATTTTGTGTGATATAATTTTATTTAGAACTTTTGTGTTT-----TAAATGGCTG 4896
Db
4856 TAGTTAAGTACTTGTGTGATATAAGGTTATTTTCAAACTATTTTGGAGTAAATGGCGT 4915
QY
4897 TTAGTCTCAGACATCTGTCAAAATTTAGCGAGTTCATTTGATCTTTTCGAAACAGGATGAGA 4956
Db
4916 TAGTAGTCAAGGACGACGTTAAGATTTCTGAGTTTCAATTTGCTGCGCTGAGAAAT 4975
QY
4957 TACTTTCCGCAATTCATGACTAAGAGTCAAGAGTGTAGAAATATCGATCTGGAACAAGATTA 5016
Db
4976 TCTTACCTGCTGTATGACTTCGGTCAAGACGCTACGAATTTTCGAAAGTTGACAAAGTGA 5035
QY
5017 TGGCTGTTAAGAAATGATAGTCTTTCTGATGTAGATTTTACTTAAAGGTGTTAAGTTAGTTA 5076
Db
5036 TTGCAATGGAAACGATTCGTTATCCGATGTGAATTTGCTTAAAGGTGTAAAGCTTTGTTA 5095
QY
5077 AGAAAGGTTATCTGTGCTTAGCTGATTTGTTAGTGTCTGCGGAGTGGAAATCTCCCGGATA 5136
Db
5096 AGGATGGTTATGTGTTTAGCAGGTTAGTTGTGTCGCGGAGTGGAACTTACCCGACA 5155
QY
5137 ACTGCGCTGGTGTGCTGAGTGTGTTGTTTGTATGTAGATAAGAGAAATGAAAGAGTGAAGAG 5196
Db
5156 ACTGCAGAGTGGAGTAGCGTTTGTGTTGTTGTTAAGAGAAATGCAAGAGATGACGAAG 5215
QY
5197 CAAAGGCTGGTGTGATATCAACGCTGCTGTTGCAAAAGAAATTTTCTTTTAAAGCTAATCC 5256
Db
5216 CAACACTTGGATCTTATAGAACCAAGTGCAGTCAAGAAACGATTTGCTTTCAAAATGATCC 5275
QY
5257 CTAAATTTTCAATACATCCGAGGATGCTGAGAGCACCCTGCGGAGGTTAGTGAATA 5316
Db
5276 CGAATTTATAGCAATTTACCCCGGATGCTGAGAGAAAGTTTGGCAAGTTTGTAGTTAATA 5335
QY
5317 TCAAAGGAGTGTGCTATGGAAGAGGATGATGCTCTTTTATCTTTTGGAGTTTCTTTCAATTT 5376
Db
5336 TTAGAGGTTGTCATGGAAGAGGTTTCTGCTCTTTATCTTTTGGAGTTTGTCTCAGTTT 5395
|||||

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 16, 2005, 21:45:12 ; Search time 1904.23 Seconds
(without alignments)
17518.955 Million cell updates/sec

Title: US-09-551-494-5
Perfect score: 6355
Sequence: 1 gatgtttaatagtttcga.....taaccgcggtagcgccca 6355

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 4134886 seqs, 2624710521 residues

Total number of hits satisfying chosen parameters: 8269772

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_23Sep04:*

- 1: Geneseq1980s:*
- 2: Geneseq1990s:*
- 3: Geneseq2000s:*
- 4: Geneseq2001as:*
- 5: Geneseq2001bs:*
- 6: Geneseq2002as:*
- 7: Geneseq2002bs:*
- 8: Geneseq2003as:*
- 9: Geneseq2003bs:*
- 10: Geneseq2003cs:*
- 11: Geneseq2003ds:*
- 12: Geneseq2004s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	6355	100.0	6355	3	AAC62372 CDNA sequ
2	6355	100.0	6355	9	Acc85005 TMV-U2 ge
3	2631	41.4	6384	2	Aaq71546 Component
4	2621.6	41.3	6384	6	AbA04211 Tomato mo
5	2468	38.8	6395	2	Aaz20642 TMV-based
6	2468	38.8	6395	4	Aaf82330 Wild-type
7	2466.4	38.8	6395	2	Aaq95155 Tobacco m
8	2466.4	38.8	6395	3	AAC62369 CDNA sequ
9	2455.4	38.6	6439	2	Aaz20643 TMV-based
10	2455.4	38.6	6439	4	Aaf82331 Tobacco m
11	2452.2	38.6	6475	2	Aaz20645 TMV-based
12	2452.2	38.6	6475	4	Aaf82333 Tobacco m
13	2449.8	38.5	6446	2	Aaz20646 TMV-based
14	2449.8	38.5	6446	4	Aaf82334 Tobacco m
15	2431.2	38.3	6425	2	Aaz20644 TMV-based
16	2431.2	38.3	6425	4	Aaf82332 Tobacco m
17	2291.4	36.1	11222	12	ADN97500 Artificial
18	2291.4	36.1	11641	6	Aad45229 TTODA (rg
19	2291.4	36.1	11641	6	Aad30539 TTODA rGA
20	2291.4	36.1	11641	10	ADD84772 Tobacco m
21	2291.4	36.1	11641	12	ADM48702 TMV based
22	2291.4	36.1	11641	12	ADM48702 TMV based

22	2289.8	36.0	10600	11	ADM68432
23	2289.8	36.0	10600	12	ADP26585
24	2289.8	36.0	10624	11	ADM68433
25	2289.8	36.0	10624	12	ADP26586
26	2289.8	36.0	11641	12	ADJ88298
27	2288.2	36.0	7685	5	AAD20298
28	2288.2	36.0	7685	5	AAD20294
29	2288.2	36.0	7685	6	AAD24478
30	2288.2	36.0	7685	6	AAD24482
31	2286.6	36.0	7685	4	AAD02010
32	2286.6	36.0	7685	5	AAD20288
33	2286.6	36.0	7685	5	AAD20291
34	2286.6	36.0	7685	6	AAD24472
35	2286.6	36.0	7685	6	AAD24475
36	2286.6	36.0	7685	10	ADB83319
37	2286.6	36.0	7685	10	ADB81243
38	2286.6	36.0	7685	10	ADB81246
39	2286.6	36.0	7686	5	AAD20289
40	2286.6	36.0	7686	6	AAD24473
41	2286.6	36.0	7686	8	ADA15012
42	2286.6	36.0	7686	10	ADB83320
43	2286.6	36.0	7686	10	ADB81244
44	2286.6	36.0	7687	5	AAD20290
45	2286.6	36.0	7687	6	AAD24474

ALIGNMENTS

RESULT 1

AAC62372
ID AAC62372 standard; DNA; 6355 BP.

XX AC AAC62372;

DT 19-MAR-2001 (first entry)

XX cDNA sequence of the genome of tobacco mosaic virus-U2.

DE Plant phenotype; gene trait; Nicotiana; Oryza sativa; Zea mays; Brassica;
KW Gossypium; Triticum; Arabidopsis; Petunia; herbicide; transgenic plant;
KW tobacco necrosis virus; TMV; tobacco mosaic virus; TMV; helper virus; ss.

OS Tobacco mosaic virus.

XX WO200063397-A2.

XX PD 26-OCT-2000.

XX PF 17-APR-2000; 2000WO-EP003521.

XX PR 20-APR-1999; 99US-00294022.

XX (AVET) AVENTIS CROPS SCIENCE NV.

XX Meulwaeter F, Cornelissen M, Jacobs J, Van Eldik G, Metzlaiff M;

XX WPI; 2000-687182/67.

XX Identifying and isolating genes involved in determining the trait or
PT phenotype of plant species, by infecting plants with gene silencing
PT constructs targeted to the gene, and identifying plants with altered
traits.

XX Example 1; Page 53-56; 64pp; English.

XX The specification describes a method for isolating genes that determine a
CC trait or phenotype of a plant species. The method comprises identifying a
CC set of nucleic acids of genes correlated with the trait, creating a
CC library of gene silencing constructs in a viral RNA vector, targeting the
CC gene silencing constructs to the nucleic acid set, infecting a collection
CC of individual plants with these, identifying plants with altered traits
CC or phenotype, and isolating genes of the invention. The method is useful

for isolating genes involved in the determination of trait or a phenotype of a plant such as Nicotiana, Oryza sativa, Zea mays, Brassica, Gossypium, Triticum, Arabidopsis or Petunia. The method is also useful for modulating the expression of selected nucleic acid sequences and for validating the function of a nucleic acid sequence whose expression is correlated with the presence or absence of a specific trait in plants, but with otherwise unknown function. The method is also useful for developing agronomically useful products such herbicides or transgenic plants. The present sequence represents the cDNA sequence of the genome of tobacco mosaic virus (TMV)-U2. The sequence was used to construct a plasmid vector for the synthesis of an infective hybrid tobacco mosaic virus (TMV)/satellite tobacco necrosis virus (STNV) helper virus RNA. This helper virus is used in the method of the invention

XX SQ Sequence 6355 BP; 1933 A; 1112 C; 1489 G; 1821 T; 0 U; 0 Other;

Query Match 100.0%; Score 6355; DB 3; Length 6355;
Best Local Similarity 100.0%; Pred. No. 0;
Matches 6355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 GATGTTTAAATAGTTTCGACAAACAACAATTAATAAACAACAAACAATTAACAACAACA 60
DB 1 GATGTTTAAATAGTTTCGACAAACAACAATTAATAAACAACAAACAATTAACAACAACA 60
QY 61 AACAAACAACATGGCACAATCAATCTATAATAGCAAGCCCTTCCTGAAAGCGTGAG 120
DB 61 AACAAACAACATGGCACAATCAATCTATAATAGCAAGCCCTTCCTGAAAGCGTGAG 120
QY 121 TGGTAAAAACATCTCGTTAAATGACCTTCGAAAGAGGCGCATGTACGATACGGCCGTGGA 180
DB 121 TGGTAAAAACATCTCGTTAAATGACCTTCGAAAGAGGCGCATGTACGATACGGCCGTGGA 180
QY 181 AGAATTAACGCCCGCGGACCGTAGACCAAAAGTCAACTTTTCCAAAACTATTAGCGAAGA 240
DB 181 AGAATTAACGCCCGCGGACCGTAGACCAAAAGTCAACTTTTCCAAAACTATTAGCGAAGA 240
QY 241 GCAAAAGCGCTTCTAGTCTCCAAAGCGGTACCGGAGTTCAGATTACCTTTTATAACTCA 300
DB 241 GCAAAAGCGCTTCTAGTCTCCAAAGCGGTACCGGAGTTCAGATTACCTTTTATAACTCA 300
QY 301 AAATGCCGTPACACAGTTTGGCTGGAGTTTGAGAGCATTAGAATTTGGAATCTGATGCT 360
DB 301 AAATGCCGTPACACAGTTTGGCTGGAGTTTGAGAGCATTAGAATTTGGAATCTGATGCT 360
QY 361 ACAAGTTCCTTATGATCGCGACATATCATATAGTGGGAACCTTTCAGACACATTTGTT 420
DB 361 ACAAGTTCCTTATGATCGCGACATATCATATAGTGGGAACCTTTCAGACACATTTGTT 420
QY 421 CAAAGGCGAGGATTAACGTGCAATGCTGATGCCCAATCTGGACATACGAGATATAATGAG 480
DB 421 CAAAGGCGAGGATTAACGTGCAATGCTGATGCCCAATCTGGACATACGAGATATAATGAG 480
QY 481 GCACGAAGCAAAAGGACTCAATTTGAGATGTAATTTGTCAGATTTGTCGTTCTTAACAA 540
DB 481 GCACGAAGCAAAAGGACTCAATTTGAGATGTAATTTGTCAGATTTGTCGTTCTTAACAA 540
QY 541 GGTAAATTCCTGAGTTTCAAGGGAGGCTTTTAAAGGTATGCAAGCTCCCAACGAGT 600
DB 541 GGTAAATTCCTGAGTTTCAAGGGAGGCTTTTAAAGGTATGCAAGCTCCCAACGAGT 600
QY 601 CTGCTGCTCTAAAACCTTTTCAAGGATGTCGAATACATCCGCGCAGAGATAGTGTAGAAG 660
DB 601 CTGCTGCTCTAAAACCTTTTCAAGGATGTCGAATACATCCGCGCAGAGATAGTGTAGAAG 660
QY 661 ATACGCTGTGCTCTGCAAGTTTGTATGATATTTCCCTGTCGATGAGTTGAGCTGCGTT 720
DB 661 ATACGCTGTGCTCTGCAAGTTTGTATGATATTTCCCTGTCGATGAGTTGAGCTGCGTT 720
QY 721 AATATCTAAGAAATATACATGATGATGACGCTTCCATTTTGGCAGAAGCATTTACT 780
DB 721 AATATCTAAGAAATATACATGATGATGATGACGCTTCCATTTTGGCAGAAGCATTTACT 780
QY 781 AGACCAGCGGATTAAGCTTAAATGAATAAGCGGCAACTTTTCAAAAGAGAGGATGATGA 840

DB 781 AGCCAGAGCGGAGTTACGCTTAATGAATAAGCGCAACTTTTCAAAAGAGAGGATGATGA 840
QY 841 TGTTCTTTTCTTTCTTTGCTGATGAAGTACTTTAAATATATAGTCAATAATACAAATAAT 900
DB 841 TGTTCTTTTCTTTCTTTGCTGATGAAGTACTTTAAATATATAGTCAATAATACAAATAAT 900
QY 901 CTGCAATATATAGTAAATCTTACTTTCCTGCTCTTAGTAGAATAGTCTTACTTAAAGGA 960
DB 901 CTGCAATATATAGTAAATCTTACTTTCCTGCTCTTAGTAGAATAGTCTTACTTAAAGGA 960
QY 961 ATTTTATAGTCACTAGGGTTAATCTTGTGTTTGTAAATTTTACCAAGTAGATACCTATAT 1020
DB 961 ATTTTATAGTCACTAGGGTTAATCTTGTGTTTGTAAATTTTACCAAGTAGATACCTATAT 1020
QY 1021 TCTGTACAAGAGTGTAGACAAGTAGGGTGTGATAGTGTAGTGTCTTATGAGGCGATGA 1080
DB 1021 TCTGTACAAGAGTGTAGACAAGTAGGGTGTGATAGTGTAGTGTCTTATGAGGCGATGA 1080
QY 1081 AGACGCTTTTCTTACAGAAACCTTGGCCATCTTCAACACTGCAAGAGCAATCTTTAG 1140
DB 1081 AGACGCTTTTCTTACAGAAACCTTGGCCATCTTCAACACTGCAAGAGCAATCTTTAG 1140
QY 1141 AGACACGCTTTCGGTTAACTTTTGGTTCCCTTAAGATGAAGGACATGCTGTAGTACCGCT 1200
DB 1141 AGACACGCTTTCGGTTAACTTTTGGTTCCCTTAAGATGAAGGACATGCTGTAGTACCGCT 1200
QY 1201 GTTTGAGGTTCTTATACAGCAAAAGATGACAAGAGTGAAGTCAATGTTAATCGTGA 1260
DB 1201 GTTTGAGGTTCTTATACAGCAAAAGATGACAAGAGTGAAGTCAATGTTAATCGTGA 1260
QY 1261 CTTTCGTTTACACAGTCTTAATCATATCAGAAATATCAAGCCCAAGCGTTTAACTTACCA 1320
DB 1261 CTTTCGTTTACACAGTCTTAATCATATCAGAAATATCAAGCCCAAGCGTTTAACTTACCA 1320
QY 1321 GAACGTTAATCTTTCGTGAGTCTATAAGATCCGCGTGTATTAATCAATGCTGTACTGC 1380
DB 1321 GAACGTTAATCTTTCGTGAGTCTATAAGATCCGCGTGTATTAATCAATGCTGTACTGC 1380
QY 1381 TAGTCTGTAAGTGGATGATGAAGCAATCTTCAACCTTGTCAATGATCTTCTTCTT 1440
DB 1381 TAGTCTGTAAGTGGATGATGAAGCAATCTTCAACCTTGTCAATGATCTTCTTCTT 1440
QY 1441 GCAGACTTAAGCTGCTGCTTCAAGACGATATAGTAATGGGAAAGTTTCGGTGTGGA 1500
DB 1441 GCAGACTTAAGCTGCTGCTTCAAGACGATATAGTAATGGGAAAGTTTCGGTGTGGA 1500
QY 1501 TAAGACCACTTCTGAACCTTATTTGGGATGAGGTGGCAAAATTTTGGGAAACGTTTTCC 1560
DB 1501 TAAGACCACTTCTGAACCTTATTTGGGATGAGGTGGGCAAAATTTTGGGAAACGTTTTCC 1560
QY 1561 CACTATCAAAGAGAGATTTGCTGAGCAGGAAATTTCTGGATGTAAAGTGAAGTCTCTGAA 1620
DB 1561 CACTATCAAAGAGAGATTTGCTGAGCAGGAAATTTCTGGATGTAAAGTGAAGTCTCTGAA 1620
QY 1621 GATCAAGATCCAGATCTGTATGTCATGTAAGAGACAGGTTCTGATGCTGAATACACAA 1680
DB 1621 GATCAAGATCCAGATCTGTATGTCATGTAAGAGACAGGTTCTGATGCTGAATACACAA 1680
QY 1681 GTCTGAGGAGTTTACCGCATCTAGATATCAAGAGGCTTAGAAGAGCTGAGCAATGTA 1740
DB 1681 GTCTGAGGAGTTTACCGCATCTAGATATCAAGAGGCTTAGAAGAGCTGAGCAATGTA 1740
QY 1741 CGACGCTTATCAGAAATTTATCTATCTTAAGGCTGCTGATTAATTTTCGATATCGCAAGTT 1800
DB 1741 CGACGCTTATCAGAAATTTATCTATCTTAAGGCTGCTGATTAATTTTCGATATCGCAAGTT 1800
QY 1801 CAAAGCATGTCGAAGGCTTTAGATGTTAGTCTGATGTCGACGACGAGTAACTGTTGC 1860
DB 1801 CAAAGCATGTCGAAGGCTTTAGATGTTAGTCTGATGTCGACGACGAGTAACTGTTGC 1860
QY 1861 AGTGGCCGAGATAGAGCGGTTTAACTTCTTACTTTTGTATAGCCCAACCGAGGAGATGT 1920

Db 1861 AGTGGCCGAGAAATAGAACGGGTTTAACTCTTACTTTTGTATAGCCAAACCGAGGAGAAATGT 1920
Qy 1921 GGCTAAGGCTCTTAAAGCAGCGGCTCTGAGGCGGTGTATGTCTTGAACCGACATCCGA 1980
Db 1921 GGCTAAGGCTCTTAAAGCAGCGGCTCTGAGGCGGTGTATGTCTTGAACCGACATCCGA 1980
Qy 1981 AGAGGTGAACGTAATAATTTTCTATTTCTGAGAGAGGAGATGGCTGTGTGTGAGA 2040
Db 1981 AGAGGTGAACGTAATAATTTTCTATTTCTGAGAGAGGAGATGGCTGTGTGTGAGA 2040
Qy 2041 AAGTCATGGTTTGACGAATGCTAACTTAGACACACAGAGTTGAGTCCCTCAACGATTT 2100
Db 2041 AAGTCATGGTTTGACGAATGCTAACTTAGACACACAGAGTTGAGTCCCTCAACGATTT 2100
Qy 2101 CCATAAGGCTTGGCTGGATAGTGTGATTTACAAAGCAATGGCATCGGTGTGCTACACTGG 2160
Db 2101 CCATAAGGCTTGGCTGGATAGTGTGATTTACAAAGCAATGGCATCGGTGTGCTACACTGG 2160
Qy 2161 CTCACTCAAGTTCAACAAATGAAGAACTATGTGGAAGCTTGGCAGCTTGGTGTGCGC 2220
Db 2161 CTCACTCAAGTTCAACAAATGAAGAACTATGTGGAAGCTTGGCAGCTTGGTGTGCGC 2220
Qy 2221 CACTGTATCAAACTCTATGCAAGTCACTAAAGATGAAGTCGGGTATGATCTGATTCAG 2280
Db 2221 CACTGTATCAAACTCTATGCAAGTCACTAAAGATGAAGTCGGGTATGATCTGATTCAG 2280
Qy 2281 GGAGAAAGTTGGTGTGGGATGTCACTTTGAAAAGTGGCTCTCAAACTCGGGCCAA 2340
Db 2281 GGAGAAAGTTGGTGTGGGATGTCACTTTGAAAAGTGGCTCTCAAACTCGGGCCAA 2340
Qy 2341 AGGTCAATTCATGGGGAGTTGCTGGATTAACAAGGGGAAATGTTTACTGCACTTCTATC 2400
Db 2341 AGGTCAATTCATGGGGAGTTGCTGGATTAACAAGGGGAAATGTTTACTGCACTTCTATC 2400
Qy 2401 TTATGAAGGATAGAAATGGTGACTGAGAGGAGCTGGAGGAGTGGCTGATCATCTGA 2460
Db 2401 TTATGAAGGATAGAAATGGTGACTGAGAGGAGCTGGAGGAGTGGCTGATCATCTGA 2460
Qy 2461 TACAATGGTATATCTGATATGCAAAATCTGAGGAAACAAATGAGAGACGG 2520
Db 2461 TACAATGGTATATCTGATATGCAAAATCTGAGGAAACAAATGAGAGACGG 2520
Qy 2521 TGAACCCACGAACTCTGCAAGATGTTGATGGATGGGTGGCTGGTGTGGAAA 2580
Db 2521 TGAACCCACGAACTCTGCAAGATGTTGATGGATGGGTGGCTGGTGTGGAAA 2580
Qy 2581 GTACAAAGGAGATTTGAAAGATTTGATCTTGATGAGGATTTGATCTTGGTCTCTGGAAA 2640
Db 2581 GTACAAAGGAGATTTGAAAGATTTGATCTTGATGAGGATTTGATCTTGGTCTCTGGAAA 2640
Qy 2641 ACAAGCTGCTCTATGATCAGAGAAAGGGCTAAATTCATCTGGAATGATAAGAGCCAAAT 2700
Db 2641 ACAAGCTGCTCTATGATCAGAGAAAGGGCTAAATTCATCTGGAATGATAAGAGCCAAAT 2700
Qy 2701 GGACAAATGTGAGAACGGTAGATTCATCTTAATGCAATCCAAACCCGGATCACCAAGAG 2760
Db 2701 GGACAAATGTGAGAACGGTAGATTCATCTTAATGCAATCCAAACCCGGATCACCAAGAG 2760
Qy 2761 GCTTTTATTCATGAGAGGTTGATGCTGCAACCGGTTGCTTAACCTTCCCTGGTGTAT 2820
Db 2761 GCTTTTATTCATGAGAGGTTGATGCTGCAACCGGTTGCTTAACCTTCCCTGGTGTAT 2820
Qy 2821 CTCTGGTTGCAACATCGCATATTTACGGAGATACACAGCAGATTCCTTTCAATTAACAG 2880
Db 2821 CTCTGGTTGCAACATCGCATATTTACGGAGATACACAGCAGATTCCTTTCAATTAACAG 2880
Qy 2881 AGTTCAGAAATTTCCGATATCCCAACATTTTGAAGCTGCAAGTGGATGAAGTTGAGAT 2940
Db 2881 AGTTCAGAAATTTCCGATATCCCAACATTTTGAAGCTGCAAGTGGATGAAGTTGAGAT 2940
Qy 2941 GAGGAGGACCACTGAGATCCCAAGTGTGAAATTTTCTTACAAATCGAAGTACGA 3000
Db 2941 GAGGAGGACCACTGAGATCCCAAGTGTGAAATTTTCTTACAAATCGAAGTACGA 3000

Qy 3001 AGGAGCGGTGACAAACCACTTCAACTGTATCAACGATCGGTCTCATCTGAGATGATAGCGG 3060
Db 3001 AGGAGCGGTGACAAACCACTTCAACTGTATCAACGATCGGTCTCATCTGAGATGATAGCGG 3060
Qy 3061 TAAGGGAGTAGTAAACAGTGTCTTCCAAACCACTAAAGGGGAAATTTGTAACCTTCACTCA 3120
Db 3061 TAAGGGAGTAGTAAACAGTGTCTTCCAAACCACTAAAGGGGAAATTTGTAACCTTCACTCA 3120
Qy 3121 GGCTGATAAATTTTGAGTTAGAGGAGAAAGGGCTATAAGAAATGTGAAACCGGTTCAAGAT 3180
Db 3121 GGCTGATAAATTTTGAGTTAGAGGAGAAAGGGCTATAAGAAATGTGAAACCGGTTCAAGAT 3180
Qy 3181 CCAAGGAGAAACCTTTGAAAGATGTGCTGCTGAGATTGACCGCAACTCCACTGACTCT 3240
Db 3181 CCAAGGAGAAACCTTTGAAAGATGTGCTGCTGAGATTGACCGCAACTCCACTGACTCT 3240
Qy 3241 GATTTCCAAGTCTTCCCGCATGTTCTAGTCGCTCTGACTAGACACAAAGAGCTTCAA 3300
Db 3241 GATTTCCAAGTCTTCCCGCATGTTCTAGTCGCTCTGACTAGACACAAAGAGCTTCAA 3300
Qy 3301 ATATTACACCGTAGTGTAGATCCCTTTAGTACAGATAATTTAGTATTGCTTCTTTAAG 3360
Db 3301 ATATTACACCGTAGTGTAGATCCCTTTAGTACAGATAATTTAGTATTGCTTCTTTAAG 3360
Qy 3361 CTCTCTCTTTTGAATAATGTATGTGTAAGACGAGTAGTAGATAGCAATTTACAGATGGA 3420
Db 3361 CTCTCTCTTTTGAATAATGTATGTGTAAGACGAGTAGTAGATAGCAATTTACAGATGGA 3420
Qy 3421 TGCAAGTGTCCAAAGGTCATAATCTCTTTGTGGCAACACCTAAATCAGAGACTTTCCAGA 3480
Db 3421 TGCAAGTGTCCAAAGGTCATAATCTCTTTGTGGCAACACCTAAATCAGAGACTTTCCAGA 3480
Qy 3481 TCTACAGTCTTATTACGATGTATGCTCTCCCTCGTGTAAATAGTACTATCTTAAACAAGTATGA 3540
Db 3481 TCTACAGTCTTATTACGATGTATGCTCTCCCTCGTGTAAATAGTACTATCTTAAACAAGTATGA 3540
Qy 3541 TGCTGTTACATGAGGTTACGTGATAATAGTCTTAAATGTGAGAGATTGTGTTCTGATTT 3600
Db 3541 TGCTGTTACATGAGGTTACGTGATAATAGTCTTAAATGTGAGAGATTGTGTTCTGATTT 3600
Qy 3601 TTCCAAAGTATTTCCGATGCCAAGGAGGTGAACCATGCTAGAGCCAGTTTTCGGTAC 3660
Db 3601 TTCCAAAGTATTTCCGATGCCAAGGAGGTGAACCATGCTAGAGCCAGTTTTCGGTAC 3660
Qy 3661 CGCGCGGAAACCGCAAGGGCTGCAGGACTACTCGAAAAATCTGGTTGCAATGATTAAGAAG 3720
Db 3661 CGCGCGGAAACCGCAAGGGCTGCAGGACTACTCGAAAAATCTGGTTGCAATGATTAAGAAG 3720
Qy 3721 AAATTTCAACGCAACGACCTGACGGGAGCGATTGACATTTGAGAGACCGCATCTGTTGT 3780
Db 3721 AAATTTCAACGCAACGACCTGACGGGAGCGATTGACATTTGAGAGACCGCATCTGTTGT 3780
Qy 3781 AGTAGATAAGTTTTTTCATAGCTATTTTATTTAAAAAGAAAAATACAAAAAATATTGC 3840
Db 3781 AGTAGATAAGTTTTTTCATAGCTATTTTATTTAAAAAGAAAAATACAAAAAATATTGC 3840
Qy 3841 TGAGATGATACGCAAGGATTCATATGATGAGATGGTTGAAAAACAGGAAAGAGTACTATT 3900
Db 3841 TGAGATGATACGCAAGGATTCATATGATGAGATGGTTGAAAAACAGGAAAGAGTACTATT 3900
Qy 3901 GGACGACTTGGCTAACTACAATTTTACAGATCTCGCGGCCATCGATCGTACAGACAT 3960
Db 3901 GGACGACTTGGCTAACTACAATTTTACAGATCTCGCGGCCATCGATCGTACAGACAT 3960
Qy 3961 GATCAAGGCTCAACCAAAACAGAAATTTGACCTTTCAATTCAGAAATGAATACCTGCTCT 4020
Db 3961 GATCAAGGCTCAACCAAAACAGAAATTTGACCTTTCAATTCAGAAATGAATACCTGCTCT 4020
Qy 4021 GCAAAACAAATTTGTCTACCAATTCGAAGCAGATCAACGGTATTTTGGCCGGTTTCTCAGAGCT 4080
Db 4021 GCAAAACAAATTTGTCTACCAATTCGAAGCAGATCAACGGTATTTTGGCCGGTTTCTCAGAGCT 4080

QY	4081	TACAAGGTGCTGCTCGAGGCAATTTGATCTTAAGAAGTTCTTTCTTTACTAGGAAAC	4140
DB	4081	TACAAGGTGCTGCTCGAGGCAATTTGATCTTAAGAAGTTCTTTCTTTACTAGGAAAC	4140
QY	4141	TCCAGAACAGATTCAGAAATTTTCTCGGATCTCGACTCGCACTGCTTCTATGGATGTGT	4200
DB	4141	TCCAGAACAGATTCAGAAATTTTCTCGGATCTCGACTCGCACTGCTTCTATGGATGTGT	4200
QY	4201	AGAACTGGATATTTCTAAGTATGATAGTCCACAGAACGAGTTTCATTTGCTGTAGATA	4260
DB	4201	AGAACTGGATATTTCTAAGTATGATAGTCCACAGAACGAGTTTCATTTGCTGTAGATA	4260
QY	4261	TGAAATATGAAAGATTTGGTCTCAATGAGTTTGGCCGAAAGTGTGGAACAAGGGCA	4320
DB	4261	TGAAATATGAAAGATTTGGTCTCAATGAGTTTGGCCGAAAGTGTGGAACAAGGGCA	4320
QY	4321	CAGGAAACAACTTTGGAAGATTTACATGCTGGAATCAAGACATGCTGTGTATCAAG	4380
DB	4321	CAGGAAACAACTTTGGAAGATTTACATGCTGGAATCAAGACATGCTGTGTATCAAG	4380
QY	4381	GAAAAGCGGTGATGTGACTACTTTTCATCGGCAATCTGTTAATAAGACAGCTTGTGCG	4440
DB	4381	GAAAAGCGGTGATGTGACTACTTTTCATCGGCAATCTGTTAATAAGACAGCTTGTGCG	4440
QY	4441	TTCAATGTTACCGATCGAAAGGTCTAATAAGGTGCTTTTGTGAGACGATTCGGTTT	4500
DB	4441	TTCAATGTTACCGATCGAAAGGTCTAATAAGGTGCTTTTGTGAGACGATTCGGTTT	4500
QY	4501	GTATTTTCCAAAGGTTTGGATTTCCCTGACATTCAGTCACTGCTTAATCTCATGTGAA	4560
DB	4501	GTATTTTCCAAAGGTTTGGATTTCCCTGACATTCAGTCACTGCTTAATCTCATGTGAA	4560
QY	4561	TTTTGAGGCCAACTGTATAGAAAGAGTACGTTTACTTTTGTGAGATACATACATA	4620
DB	4561	TTTTGAGGCCAACTGTATAGAAAGAGTACGTTTACTTTTGTGAGATACATACATA	4620
QY	4621	CCATGATAGGAGCAATAGTGTATATGATCCCTTGAAGTTGATCTCCAACTTGGGG	4680
DB	4621	CCATGATAGGAGCAATAGTGTATATGATCCCTTGAAGTTGATCTCCAACTTGGGG	4680
QY	4681	AAAAACATATCAAGGATTTATGATCACTTAGAAGAGTTAAGGTTGCTTTTGTGCGATGTC	4740
DB	4681	AAAAACATATCAAGGATTTATGATCACTTAGAAGAGTTAAGGTTGCTTTTGTGCGATGTC	4740
QY	4741	TTGTTCCGCTCGAACTGGTGTAGGCTTTCCGAGCTGAAACGAGCTATCAAGGAGT	4800
DB	4741	TTGTTCCGCTCGAACTGGTGTAGGCTTTCCGAGCTGAAACGAGCTATCAAGGAGT	4800
QY	4801	TCATAAAACCGGATTTGATGTTGCTTTGCTTTTAAATGTTTAACTAACAAATTTTGTGTA	4860
DB	4801	TCATAAAACCGGATTTGATGTTGCTTTGCTTTTAAATGTTTAACTAACAAATTTTGTGTA	4860
QY	4861	TAAATTTTATTTAGAACTTTGTTTAAATGGCTTTAGTCTCAGAGATACTGTCAAAA	4920
DB	4861	TAAATTTTATTTAGAACTTTGTTTAAATGGCTTTAGTCTCAGAGATACTGTCAAAA	4920
QY	4921	TTAGCGAGTCTATTCATCTTTGAAACAGGATGAGATCTTCCGGCATTCATGCTAAGG	4980
DB	4921	TTAGCGAGTCTATTCATCTTTGAAACAGGATGAGATCTTCCGGCATTCATGCTAAGG	4980
QY	4981	TCAAGAGTGTAGAAATATCGACTGTGGAACAAGATTAAGGCTGTTAAGAAATGATAGTCTTT	5040
DB	4981	TCAAGAGTGTAGAAATATCGACTGTGGAACAAGATTAAGGCTGTTAAGAAATGATAGTCTTT	5040
QY	5041	CTGATGTAGATTTACTTAAAGGTGTAAAGTGTAGTTAAGAAAGGTATGTGTCTTAGCTG	5100
DB	5041	CTGATGTAGATTTACTTAAAGGTGTAAAGTGTAGTTAAGAAAGGTATGTGTCTTAGCTG	5100
QY	5101	ATTTGGTGTGCTGCGGAGTGGAACTCTCCCGGATTAATCTCCGCTGGTGTGCTAGTGT	5160
DB	5101	ATTTGGTGTGCTGCGGAGTGGAACTCTCCCGGATTAATCTCCGCTGGTGTGCTAGTGT	5160
QY	5161	GTATTTGTAGATAAGAGAAATGAAAGGATAGGAAGCAACGCTGGGTGCTATCACGCC	5220

DB	5161	GTATTTGTAGATAAGAGAAATGAAAGGAGTAAAGGACCAACGCTGGGTGCTATCACGCC	5220
QY	5221	CTGCTTGCAGAAAGAAATTTTCTTTTAAGCTAAATCCCTAAATTTATTTCAATAATCCGAGG	5280
DB	5221	CTGCTTGCAGAAAGAAATTTTCTTTTAAGCTAAATCCCTAAATTTATTTCAATAATCCGAGG	5280
QY	5281	ATGCTGAGAGAACCCGCTGGCAAGTGTAGTGAATATCAAAAGGAGTGCCTATGGAAGAAG	5340
DB	5281	ATGCTGAGAGAACCCGCTGGCAAGTGTAGTGAATATCAAAAGGAGTGCCTATGGAAGAAG	5340
QY	5341	GATACGTGCTTTATCTTTTGGAGTTCGTTTCAATTTGTGTAGTACATATAAATAATGTAA	5400
DB	5341	GATACGTGCTTTATCTTTTGGAGTTCGTTTCAATTTGTGTAGTACATATAAATAATGTAA	5400
QY	5401	GAAAAGTTTCAGGGAACGTAATTTTGAGTGTGACAGACGGCTCGCCAAATTTGAACTCACTG	5460
DB	5401	GAAAAGTTTCAGGGAACGTAATTTTGAGTGTGACAGACGGCTCGCCAAATTTGAACTCACTG	5460
QY	5461	AAAAGGTTTGTGAGGAGTTCGTGATGAAGTACCAATGGCTGTGGAATCTCGAAAGGTTTC	5520
DB	5461	AAAAGGTTTGTGAGGAGTTCGTGATGAAGTACCAATGGCTGTGGAATCTCGAAAGGTTTC	5520
QY	5521	CGGAAAACAAAAAGAAATGCTAGCTAAATTAATGTTTAATAAAGAAATAAATAACAGTG	5580
DB	5521	CGGAAAACAAAAAGAAATGCTAGCTAAATTAATGTTTAATAAAGAAATAAATAACAGTG	5580
QY	5581	GTAAGAGGGTTTAAATTTGAGGAAATGAGGAAATTAAGGATTAAGTGAATGAGTCTATCG	5640
DB	5581	GTAAGAGGGTTTAAATTTGAGGAAATGAGGAAATTAAGGATTAAGTGAATGAGTCTATCG	5640
QY	5641	CGTCATCGAGTACGTTTAAATTAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTT	5700
DB	5641	CGTCATCGAGTACGTTTAAATTAATGCTTAAATGCTTAAATGCTTAAATGCTTAAATGCTT	5700
QY	5701	TTACTTATCTTCCGCTTTACGAGATCTGTCAGCTGATCAATCTCTGTGTAACAAATGCTT	5760
DB	5701	TTACTTATCTTCCGCTTTACGAGATCTGTCAGCTGATCAATCTCTGTGTAACAAATGCTT	5760
QY	5761	GGGTAAACAGTTTCAACAGCAACAGCTAGGACCAACAGTCCAAACAGCAATTTCCGAGTGC	5820
DB	5761	GGGTAAACAGTTTCAACAGCAACAGCTAGGACCAACAGTCCAAACAGCAATTTCCGAGTGC	5820
QY	5821	CTGAAAACCTGTGCTAGTATGACAGTGAATTTCTGCAATCGATTTCTTATGTTGATAG	5880
DB	5821	CTGAAAACCTGTGCTAGTATGACAGTGAATTTCTGCAATCGATTTCTTATGTTGATAG	5880
QY	5881	ATATAATTCAGCGTTGATCCGTTGATCAACGCGGTTTAAATAAGCTTTGATAGTAA	5940
DB	5881	ATATAATTCAGCGTTGATCCGTTGATCAACGCGGTTTAAATAAGCTTTGATAGTAA	5940
QY	5941	TAGAAATAAGAGTTTGAATAATCAACCCGACCGAATCTACTGAAATCGTTAAACCGGAC	6000
DB	5941	TAGAAATAAGAGTTTGAATAATCAACCCGACCGAATCTACTGAAATCGTTAAACCGGAC	6000
QY	6001	TCAGAGGTAGACGATGCTACTGTAGCTATAAGGCTTCAATCAATAATTTGGCTAATGA	6060
DB	6001	TCAGAGGTAGACGATGCTACTGTAGCTATAAGGCTTCAATCAATAATTTGGCTAATGA	6060
QY	6061	ACTGTTTCGTTGGAACCTGGCATGTTCAATCAAGCAGCTTTGAGACTGTAGTGGACTTGT	6120
DB	6061	ACTGTTTCGTTGGAACCTGGCATGTTCAATCAAGCAGCTTTGAGACTGTAGTGGACTTGT	6120
QY	6121	CTGGACCAACAACTCCGGCTACTTAGCTATGTTGTGAGATTTTCTTAAATAAAGTCCGCTG	6180
DB	6121	CTGGACCAACAACTCCGGCTACTTAGCTATGTTGTGAGATTTTCTTAAATAAAGTCCGCTG	6180
QY	6181	AGAATTAAATTCAGGGTGGCTGATACCAAAATCAGCAGTGGTGTGCTCCACTTAA	6240
DB	6181	AGAATTAAATTCAGGGTGGCTGATACCAAAATCAGCAGTGGTGTGCTCCACTTAA	6240
QY	6241	TATAACGATTTGATATCTGGAATCCAAAGTTTAAACCAATGATGCTGTATCTGCTGTA	6300

Db 6241 TATAACGATTGTCTATCTCGATCAACAGTTAAACCAATGTGATGGTGTATCTGTGTA 6300
Qy 6301 TGGCGTAAACATCGGAGAGTTTGAATCTCCCTTAAACCGCGGTAGCGGCCA 6355
Db 6301 TGGCGTAAACATCGGAGAGTTTGAATCTCCCTTAAACCGCGGTAGCGGCCA 6355
RESULT 2
ID ACC85005 standard; DNA; 6355 BP.
AC ACC85005;
XX 13-OCT-2003 (first entry)
DT TMV-U2 genome nucleotide sequence.
DE TMV-U2 genome nucleotide sequence.
XX Inhibitory RNA; viral RNA vector; coat protein; TMV; U2; gene; ds.
XX Tobacco mosaic virus.
OS WO2003052108-A2.
XX 26-JUN-2003.
XX 05-DEC-2002; 2002WO-EP013964.
XX 18-DEC-2001; 2001US-0340488P.
XX (PABB) BAYER BIOSCIENCE NV.
XX Metzlaiff MH, Gossele VML, Meulewaeter F, Fache ICA;
XX WPI; 2003-523529/49.
XX Introducing inhibitory RNA into a plant cell comprises providing a viral
XX RNA vector derived from a satellite RNA virus that encodes a coat
XX protein, and infecting a plant with the viral RNA vector and a
XX corresponding helper virus.
XX Example; Page 79-82; 86pp; English.
XX The invention relates to introducing inhibitory RNA into a plant cell.
XX The method involves providing a viral RNA vector derived from a satellite
XX RNA virus having a sequence that encodes a coat protein, and infecting a
XX plant with the viral RNA vector and a corresponding helper virus. The
XX methods and viral RNA vectors are useful in introducing inhibitory RNA
XX into plant cells. These may be used to determine or validate the function
XX of isolated nucleic acid sequences in plants. The present sequence
XX represents the nucleotide sequence of the genome of tobacco mosaic virus
XX (TMV) -U2
SQ Sequence 6355 BP; 1933 A; 1112 C; 1489 G; 1821 T; 0 U; 0 Other;
Query Match 100.08; Score 6355; DB 9; Length 6355;
Best Local Similarity 100.08; Pred. No. 0;
Matches 6355; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 1 GATGTTTTTAATAGTTTTCGACAAACAATTTAAACAAACAAACATATTACAAACACA 60
Db 1 GATGTTTTTAATAGTTTTCGACAAACAATTTAAACAAACAAACATATTACAAACACA 60
Qy 61 AACAAACAATGGCACAATACATCTATTAATAGCAACGCCCTTCTTTGAAAGCGTGAG 120
Db 61 AACAAACAATGGCACAATACATCTATTAATAGCAACGCCCTTCTTTGAAAGCGTGAG 120
Qy 121 TGGTAAACAACTCTCGTTAATGACCTTGCAGAAAGGCGCATGTACGATACGGCGGTGGA 180
Db 121 TGGTAAACAACTCTCGTTAATGACCTTGCAGAAAGGCGCATGTACGATACGGCGGTGGA 180
Qy 181 AGAATTTAAACCGCGGCGGTAGCAACCAAGTCAACTTTTCCAAACATATTAGCGGAAGA 240
Db 181 AGAATTTAAACCGCGGCGGTAGCAACCAAGTCAACTTTTCCAAACATATTAGCGGAAGA 240

Qy 241 GCAAACCGTTCTAGTCTCCAAACGGTATCCCGAGTTCAGAGATTACCTTTTATAAFACTCA 300
Db 241 GCAAACCGTTCTAGTCTCCAAACGGTATCCCGAGTTCAGAGATTACCTTTTATAAFACTCA 300
Qy 301 AAATGCCGTACACAGTTTGGCTGGAGTTTGGAGAGCATTTAGAAATTTGGAATATCTGATGCT 360
Db 301 AAATGCCGTACACAGTTTGGCTGGAGTTTGGAGAGCATTTAGAAATTTGGAATATCTGATGCT 360
Qy 361 ACAAGTTCCCTATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 420
Db 361 ACAAGTTCCCTATATGATATGATATGATATGATATGATATGATATGATATGATATGATATGAT 420
Qy 421 CAAAGGCAAGGATTTACGTGCAATTCCTGTAATGCAATTCCTGTAATGCAATTCCTGTAATGCAAT 480
Db 421 CAAAGGCAAGGATTTACGTGCAATTCCTGTAATGCAATTCCTGTAATGCAATTCCTGTAATGCAAT 480
Qy 481 GCAGAAAGGACAAAGGACTCAATTTGAGATGATATTTGTCAGATGATATTTGTCAGATGATATTTGTCAGAT 540
Db 481 GCAGAAAGGACAAAGGACTCAATTTGAGATGATATTTGTCAGATGATATTTGTCAGATGATATTTGTCAGAT 540
Qy 541 GGTAAATTCCTGAGTTTCAAAGGGAGGCTTTTAAACAGGTATGCAGAGCTCCCAACGAAGT 600
Db 541 GGTAAATTCCTGAGTTTCAAAGGGAGGCTTTTAAACAGGTATGCAGAGCTCCCAACGAAGT 600
Qy 601 CTGCTGCTCTAAACCTTTTCAAGGATTTGTCGAATACATCCCGCAGAGATAGTGTAGAG 660
Db 601 CTGCTGCTCTAAACCTTTTCAAGGATTTGTCGAATACATCCCGCAGAGATAGTGTAGAG 660
Qy 661 ATACGCTGTTGCTCTGCACAGTTTGTATGATATTTCTGTGATGATGATGATGATGATGATGATGATGAT 720
Db 661 ATACGCTGTTGCTCTGCACAGTTTGTATGATATTTCTGTGATGATGATGATGATGATGATGATGATGAT 720
Qy 721 AATATCTAAGAAATATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Db 721 AATATCTAAGAAATATACATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 780
Qy 781 AGACGACGAGGATTTAGCTTAATGAAATAGGCGCAACTTTCAAAGAGAGAGTGA 840
Db 781 AGACGACGAGGATTTAGCTTAATGAAATAGGCGCAACTTTCAAAGAGAGAGTGA 840
Qy 841 TGTGTTCTTTTCTGCTGATGAAAGTACTTTTAAATTTATAGTCTATAAATACAAAATAT 900
Db 841 TGTGTTCTTTTCTGCTGATGAAAGTACTTTTAAATTTATAGTCTATAAATACAAAATAT 900
Qy 901 CTTGCATTATGATGATTAATCTTCTTCTGCTCTCTAGTAGAATAGTTTACTTTAAGGA 960
Db 901 CTTGCATTATGATGATTAATCTTCTTCTGCTCTCTAGTAGAATAGTTTACTTTAAGGA 960
Qy 961 ATTTTGTAGTCTAGGTTTAAATCTTGGTTTGTAAATTTACCAAGTAGATACCTATAT 1020
Db 961 ATTTTGTAGTCTAGGTTTAAATCTTGGTTTGTAAATTTTACCAAGTAGATACCTATAT 1020
Qy 1021 TCTGTACAAGAGTGTAGACAAGTAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Db 1021 TCTGTACAAGAGTGTAGACAAGTAGGTGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1080
Qy 1081 AGACGCTTTGCTTTACAAGAAACCTTTGGCCATGTTCAACACTGAAGAGCAATCTTTAG 1140
Db 1081 AGACGCTTTGCTTTACAAGAAACCTTTGGCCATGTTCAACACTGAAGAGCAATCTTTAG 1140
Qy 1141 AGACGCTTTGCTTTAACTTTTGGTTTCCCTTAAAGTAGGAGCATGTTGATGATGATGATGATGATGATGAT 1200
Db 1141 AGACGCTTTGCTTTAACTTTTGGTTTCCCTTAAAGTAGGAGCATGTTGATGATGATGATGATGATGATGAT 1200
Qy 1201 GTTTGAGGGTCTTATTTACCAAGAAAGATGCAAGGAGTGAAGTCAATTTGTTAATCGTGA 1260
Db 1201 GTTTGAGGGTCTTATTTACCAAGAAAGATGCAAGGAGTGAAGTCAATTTGTTAATCGTGA 1260
Qy 1261 CTTGTTTACACAGTGTCTTATCATATCAGAACATATCAAGCCAAAGCGTTAACTTACCA 1320
Db 1261 CTTGTTTACACAGTGTCTTATCATATCAGAACATATCAAGCCAAAGCGTTAACTTACCA 1320

QY 1321 GAAAGTATTATCTTTCTGTTGAGTCTATAAGATCCCGCGTGATTAATCAATGGTGTACTGC 1380
DB 1321 GAAAGTATTATCTTTCTGTTGAGTCTATAAGATCCCGCGTGATTAATCAATGGTGTACTGC 1380
QY 1381 TAGGTCCTGAATGGAGTATAGATAAGCAATCTTCAACCTTGTCAATGACATCTTCTTT 1440
DB 1381 TAGGTCCTGAATGGAGTATAGATAAGCAATCTTCAACCTTGTCAATGACATCTTCTTT 1440
QY 1441 GCAGACTAAGCTGGCTGCGCTTCAAGACGATATAGTAATGGGAAGTCTTGGTGTCTGGA 1500
DB 1441 GCAGACTAAGCTGGCTGCGCTTCAAGACGATATAGTAATGGGAAGTCTTGGTGTCTGGA 1500
QY 1501 TAAGACCACTTCGAACTTATTTGGGATGAGTGGGCAAAATTTTGGAAAAGTCTTCCC 1560
DB 1501 TAAGACCACTTCGAACTTATTTGGGATGAGTGGGCAAAATTTTGGAAAAGTCTTCCC 1560
QY 1561 CACTATCAAGAGAGATTCGTGAGCAGGAAATCTCGGATGTAAAGTGAATGCTCTGAA 1620
DB 1561 CACTATCAAGAGAGATTCGTGAGCAGGAAATCTCGGATGTAAAGTGAATGCTCTGAA 1620
QY 1621 GATCAAGATCCCAGATCTGTATGTCCATGGAAGACAGGTTCTGTAGCTGAATACACAA 1680
DB 1621 GATCAAGATCCCAGATCTGTATGTCCATGGAAGACAGGTTCTGTAGCTGAATACACAA 1680
QY 1681 GTCTGAGGAGTTACCGCATCTAGATATCAAGAGGACTTTAGAGAGCTGAGCAATGTA 1740
DB 1681 GTCTGAGGAGTTACCGCATCTAGATATCAAGAGGACTTTAGAGAGCTGAGCAATGTA 1740
QY 1741 CGACGCGTTATCAGAAATATCTATCTTAAAGGCTGTGATTAATTTTGGATATCGCAAGTT 1800
DB 1741 CGACGCGTTATCAGAAATATCTATCTTAAAGGCTGTGATTAATTTTGGATATCGCAAGTT 1800
QY 1801 CAAAGCATGTGCAAGGCTTTAGATGTCTGATGTGGCAGCACGAGTAATCGTTGC 1860
DB 1801 CAAAGCATGTGCAAGGCTTTAGATGTCTGATGTGGCAGCACGAGTAATCGTTGC 1860
QY 1861 AGTGCCGAGATAGAGCGGTTTAACTCTTACTTTTGAATAGCCAAACCGAGGAGATGT 1920
DB 1861 AGTGCCGAGATAGAGCGGTTTAACTCTTACTTTTGAATAGCCAAACCGAGGAGATGT 1920
QY 1921 GGTAAAGCTCTTAAAGCAGCGCTCTGAGCGCTGTGTATGTCTTGAACCGACATCCGA 1980
DB 1921 GGTAAAGCTCTTAAAGCAGCGCTCTGAGCGCTGTGTATGTCTTGAACCGACATCCGA 1980
QY 1981 AGAGTGAACTGAATAAATTTCTTATGCTGAGAAAGGAGATGCTGTGTGTGCGAGA 2040
DB 1981 AGAGTGAACTGAATAAATTTCTTATGCTGAGAAAGGAGATGCTGTGTGTGCGAGA 2040
QY 2041 AAGTCATGTTTGACGAATGCTAACTTAGACACCGAGGTTGGAGTCCCTCAACGATTT 2100
DB 2041 AAGTCATGTTTGACGAATGCTAACTTAGACACCGAGGTTGGAGTCCCTCAACGATTT 2100
QY 2101 CCATAGGCTTTCGTTGAGTGTGATTAAGCAAAATGGCATCGGTTGCTACACTGG 2160
DB 2101 CCATAGGCTTTCGTTGAGTGTGATTAAGCAAAATGGCATCGGTTGCTACACTGG 2160
QY 2161 CTCACCTAAAGTTTCAAAGTGAATCTATGTGGAAGATTTGGCAGCTTGGTGTGCGC 2220
DB 2161 CTCACCTAAAGTTTCAAAGTGAATCTATGTGGAAGATTTGGCAGCTTGGTGTGCGC 2220
QY 2221 CACTGTATCAATCTATGCAAGTCACTAAAGATGAAGTCCGGTATGATTTCTGATCCAG 2280
DB 2221 CACTGTATCAATCTATGCAAGTCACTAAAGATGAAGTCCGGTATGATTTCTGATCCAG 2280
QY 2281 GGAGAAAGTGTGGTGTGGGATGCTCTTGAAGAGTGGCTCTCTCAACCTCGGGGCAA 2340
DB 2281 GGAGAAAGTGTGGTGTGGGATGCTCTTGAAGAGTGGCTCTCTCAACCTCGGGGCAA 2340
QY 2341 AGGTCATTCTAGGGAGTTGCTCGGATTAAGGGGAAATGTTTACTGACATCTTATC 2400
DB 2341 AGGTCATTCTAGGGAGTTGCTCGGATTAAGGGGAAATGTTTACTGACATCTTATC 2400
QY 2401 TTATGAGGAGATAGAAATGGTCACTGAGACGCACTGGAGGAGGCTGTGCTATCATCTGA 2460

DB 2401 TTATGAGGAGATAGAAATGGTCACTGAGACGCACTGGAGGAGGCTGTGCTATCATCTGA 2460
QY 2461 TACAATGTTATATCTTCTGATTTGAAAGCTCCAAATCTGAGGAAACAAATGAGAGACGG 2520
DB 2461 TACAATGTTATATCTTCTGATTTGAAAGCTCCAAATCTGAGGAAACAAATGAGAGACGG 2520
QY 2521 TGAACCCACCAACCTCTGCAAGATGTTGTTGATGAGGATTTGATCTTGGTGTGGAAA 2580
DB 2521 TGAACCCACCAACCTCTGCAAGATGTTGTTGATGAGGATTTGATCTTGGTGTGGAAA 2580
QY 2581 GTCAAAAGGAGATTTTGAAGATTTGATCTTGAATGAGGATTTGATCTTGGTGTGGAAA 2640
DB 2581 GTCAAAAGGAGATTTTGAAGATTTGATCTTGAATGAGGATTTGATCTTGGTGTGGAAA 2640
QY 2641 ACAAGCTCTCTGATGATCAGAGAGGCTTAATTTCAATCTGAGCTGATGAGAGCCACAAT 2700
DB 2641 ACAAGCTCTCTGATGATCAGAGAGGCTTAATTTCAATCTGAGCTGATGAGAGCCACAAT 2700
QY 2701 GGACAACTGTGAGAACGGTGTAGATTTCACTTCAATCCAAACCGCGATCACACAAG 2760
DB 2701 GGACAACTGTGAGAACGGTGTAGATTTCACTTCAATCCAAACCGCGATCACACAAG 2760
QY 2761 GCTTTTATGATGAAGGTTGATGCTGCACACACCGGTTGTGTTAACTTCTTGGTGTCTTAT 2820
DB 2761 GCTTTTATGATGAAGGTTGATGCTGCACACACCGGTTGTGTTAACTTCTTGGTGTCTTAT 2820
QY 2821 CTCTGGTTCGACATCCATACATTTAGGAGATACACACGAGATTCCTTTCATTAACAG 2880
DB 2821 CTCTGGTTCGACATCCATACATTTAGGAGATACACACGAGATTCCTTTCATTAACAG 2880
QY 2881 AGTTTCAAGATTTCCGCTATCCCAACATTTTGAAGCTGCAAGTGGATGAAGTTGAGAT 2940
DB 2881 AGTTTCAAGATTTCCGCTATCCCAACATTTTGAAGCTGCAAGTGGATGAAGTTGAGAT 2940
QY 2941 GAGGAGACCACTGAGATGCTCCAGGTTGATGTTTCTCAATGCAAGTACGA 3000
DB 2941 GAGGAGACCACTGAGATGCTCCAGGTTGATGTTTCTCAATGCAAGTACGA 3000
QY 3001 AGGAGCGGTGACACCACTTCACTGATACACGATCGCTCTCATCTGAGATGAGGCGG 3060
DB 3001 AGGAGCGGTGACACCACTTCACTGATACACGATCGCTCTCATCTGAGATGAGGCGG 3060
QY 3061 TAAGGAGTACTAAACAGTGTTCCTCAACCACTTAAAGGGAATTTGATCTTCACTCA 3120
DB 3061 TAAGGAGTACTAAACAGTGTTCCTCAACCACTTAAAGGGAATTTGATCTTCACTCA 3120
QY 3121 GGCTGATAAATTTGAGTTAGGAGAGGCTATAGAAATGTGAACACCGTTCAATGAGAT 3180
DB 3121 GGCTGATAAATTTGAGTTAGGAGAGGCTATAGAAATGTGAACACCGTTCAATGAGAT 3180
QY 3181 CCAAGGAGAAACCTTTGAAGATGTGCTGGTCAAGATTCACGCAACTCCACTGACTCT 3240
DB 3181 CCAAGGAGAAACCTTTGAAGATGTGCTGGTCAAGATTCACGCAACTCCACTGACTCT 3240
QY 3241 GATTTCCAGTCTTCCCGCATGTTCTAGTCTGCTGACTAGACACAAAGAGCTTCAA 3300
DB 3241 GATTTCCAGTCTTCCCGCATGTTCTAGTCTGCTGACTAGACACAAAGAGCTTCAA 3300
QY 3301 ATATTACACCGTAGTGTAGATCTTGTAGTACAGATTAATTTAGTGTCTTCTTTAAG 3360
DB 3301 ATATTACACCGTAGTGTAGATCTTGTAGTACAGATTAATTTAGTGTCTTCTTTAAG 3360
QY 3361 CTCCTCTCTTTTGAAGATGTATGTTAGAGCAGGCTAGTAGATAGCAATTAAGATGGA 3420
DB 3361 CTCCTCTCTTTTGAAGATGTATGTTAGAGCAGGCTAGTAGATAGCAATTAAGATGGA 3420
QY 3421 TGCAAGTGTCAAAGTCAATCTCTTTGTGGCAACCTTAAATCAGGAGACTTTCAGA 3480
DB 3421 TGCAAGTGTCAAAGTCAATCTCTTTGTGGCAACCTTAAATCAGGAGACTTTCAGA 3480
QY 3481 TCTACAGTCTTCTATGATGATGCTCCCTGCTGATAGTACTTACTTAAACAGTATGA 3540

Db 3481 TCTACAGTTCATTACAGATGATGCTCCCTGGTAAATAGTACTATATCTTAACAAGTATGA 3540
Qy 3541 TGCCTGTTACCATGAGGTTACGTAATAGTCTTAAATGTGAAGGATGTGCTCTGATTT 3600
Db 3541 TGCCTGTTACCATGAGGTTACGTAATAGTCTTAAATGTGAAGGATGTGCTCTGATTT 3600
Qy 3601 TTCCAAAAGTATTCGGATGCCAAGAGGAGTGAACCATGTCTAGAGCCAGTTTTGCCGTAC 3660
Db 3601 TTCCAAAAGTATTCGGATGCCAAGAGGAGTGAACCATGTCTAGAGCCAGTTTTGCCGTAC 3660
Qy 3661 CGCGCGGAAACCGCCAAAGGCTGACGAGTACTCTGAAAATCTGGTTGCAATGATTAAGAAG 3720
Db 3661 CGCGCGGAAACCGCCAAAGGCTGACGAGTACTCTGAAAATCTGGTTGCAATGATTAAGAAG 3720
Qy 3721 AAAATTTCAACGACACAGACCTGACGGGACGATGACATTCAGAGCACCCGATCTGTTGT 3780
Db 3721 AAAATTTCAACGACACAGACCTGACGGGACGATGACATTCAGAGCACCCGATCTGTTGT 3780
Qy 3781 AGTAGATAAGTTTTTGTAGTACTATTTTATTAAGAAAGAAAATACACAAAATATATGCG 3840
Db 3781 AGTAGATAAGTTTTTGTAGTACTATTTTATTAAGAAAGAAAATACACAAAATATATGCG 3840
Qy 3841 TGGAGTGATGACGAAGGATTCGAATGATGAGATGGTTGGAAAACAGGAAGAGTACTATT 3900
Db 3841 TGGAGTGATGACGAAGGATTCGAATGATGAGATGGTTGGAAAACAGGAAGAGTACTATT 3900
Qy 3901 GGACGACTTGGCTAACTAATTTTACAGATCTCGCGGCCATCGATCAGTACAAAGCACAT 3960
Db 3901 GGACGACTTGGCTAACTAATTTTACAGATCTCGCGGCCATCGATCAGTACAAAGCACAT 3960
Qy 3961 GATCAGGCTCAACCAAAACAGAAATTCGACCTTTCATTCAGATGAAATACCCCTGCTCT 4020
Db 3961 GATCAGGCTCAACCAAAACAGAAATTCGACCTTTCATTCAGATGAAATACCCCTGCTCT 4020
Qy 4021 GCAAAACATTTGTCTACCAATTCGAGCAGATCAACGGTATTTTGGCCGGTTTCTCAGAGCT 4080
Db 4021 GCAAAACATTTGTCTACCAATTCGAGCAGATCAACGGTATTTTGGCCGGTTTCTCAGAGCT 4080
Qy 4081 TACAAGGTTGCTGCTCGAGGCAATTTGATTTCAAGAAAGTTCTCTTTTCTTACTAGGAAAC 4140
Db 4081 TACAAGGTTGCTGCTCGAGGCAATTTGATTTCAAGAAAGTTCTCTTTTCTTACTAGGAAAC 4140
Qy 4141 TCCAGAACAGATTCAGAAATTTTCTCGGATCTCGACTCGACGTTCTCTATGGATGTGTT 4200
Db 4141 TCCAGAACAGATTCAGAAATTTTCTCGGATCTCGACTCGACGTTCTCTATGGATGTGTT 4200
Qy 4201 AGAATCGATATTTCTAAGTATGATAAGTCAAGAAACGAGTTTCAATTCGTCTGTAGAGTA 4260
Db 4201 AGAATCGATATTTCTAAGTATGATAAGTCAAGAAACGAGTTTCAATTCGTCTGTAGAGTA 4260
Qy 4261 TGAATATGGAAGATTTGGGTCTCAATGAGTTTTTGGCCGAAGTGTGGAAACAAAGGCA 4320
Db 4261 TGAATATGGAAGATTTGGGTCTCAATGAGTTTTTGGCCGAAGTGTGGAAACAAAGGCA 4320
Qy 4321 CAGGAAACAACTTTGAAAGGATTAATTCGTGGAAATCAAGACATGCTGTGGTATCAAAG 4380
Db 4321 CAGGAAACAACTTTGAAAGGATTAATTCGTGGAAATCAAGACATGCTGTGGTATCAAAG 4380
Qy 4381 GAAAGCCGTGATGTGACTTTTCAATCGGCAATCTGTTAATAGCAGCTTGTCTGGG 4440
Db 4381 GAAAGCCGTGATGTGACTTTTCAATCGGCAATCTGTTAATAGCAGCTTGTCTGGG 4440
Qy 4441 TTCAATGTTACCGATGGAAAAGGTCATAAAAGGTCTTTTGTGGAGACGATTCGGTTTT 4500
Db 4441 TTCAATGTTACCGATGGAAAAGGTCATAAAAGGTCTTTTGTGGAGACGATTCGGTTTT 4500
Qy 4501 GTATTTTCCAAAGGGTTTGGATTTCCCTGACATTCAGTCAATGCTGATCTCATGTGGAA 4560
Db 4501 GTATTTTCCAAAGGGTTTGGATTTCCCTGACATTCAGTCAATGCTGATCTCATGTGGAA 4560
Qy 4561 TTTTCAGGCCAACTGTATAGAAAGAGGTACGGTTTACTTTTGTGGTAGATACATACACA 4620
Db 4561 TTTTCAGGCCAACTGTATAGAAAGAGGTACGGTTTACTTTTGTGGTAGATACATACACA 4620

Qy 4621 CCATGATAAGGGAGCAATAGTGTATATGATCCTTTGAAGTTGATCTCCAAACTTGGGCG 4680
Db 4621 CCATGATAAGGGAGCAATAGTGTATATGATCCTTTGAAGTTGATCTCCAAACTTGGGCG 4680
Qy 4681 AAAACATATCAAGGATTTATGATCACTTAGAAGAGTTAAGGGTGTCTTTGTGCGATGTTGC 4740
Db 4681 AAAACATATCAAGGATTTATGATCACTTAGAAGAGTTAAGGGTGTCTTTGTGCGATGTTGC 4740
Qy 4741 TTGTTCCGCTCGGAAACTGGTCTTTAGGCTTTCCGACGTGAACGAGCTATCAAGGAGGT 4800
Db 4741 TTGTTCCGCTCGGAAACTGGTCTTTAGGCTTTCCGACGTGAACGAGCTATCAAGGAGGT 4800
Qy 4801 TCATAAAAACCGCGATTTGATGTTGCTTTTAAATTTGTTTAACTTTTAACTTTTGTGTGA 4860
Db 4801 TCATAAAAACCGCGATTTGATGTTGCTTTTAAATTTGTTTAACTTTTAACTTTTGTGTGA 4860
Qy 4861 TAAATTTTAAATTTTAACTTTTAAATTTGTTTAAATTTGTTTAACTTTTAACTTTTGTGTGA 4920
Db 4861 TAAATTTTAAATTTTAACTTTTAAATTTGTTTAAATTTGTTTAACTTTTAACTTTTGTGTGA 4920
Qy 4921 TTACGAGTTCATTTGATCTTTTCGAAACAGGATGAGATCTTTCCGCGCATTCATGACTAAGG 4980
Db 4921 TTACGAGTTCATTTGATCTTTTCGAAACAGGATGAGATCTTTCCGCGCATTCATGACTAAGG 4980
Qy 4981 TCAAGAGTGTAGAAATATCGACTGTGGACAAAGATTAAGGCTGTTAAGAAATGATAGTCTTT 5040
Db 4981 TCAAGAGTGTAGAAATATCGACTGTGGACAAAGATTAAGGCTGTTAAGAAATGATAGTCTTT 5040
Qy 5041 CTGATGTAGATTTTAAAGTGTAAAGTGTAAAGTGTAAAGAAAGGATGATGCTTTAGTCTG 5100
Db 5041 CTGATGTAGATTTTAAAGTGTAAAGTGTAAAGTGTAAAGAAAGGATGATGCTTTAGTCTG 5100
Qy 5101 ATTTGGTAGTCTCGGGAGTGGAAATCTCCCGGATACTCCCGTGGTGGTGTCTAGTGTCTT 5160
Db 5101 ATTTGGTAGTCTCGGGAGTGGAAATCTCCCGGATACTCCCGTGGTGGTGTCTAGTGTCTT 5160
Qy 5161 GTATTGTAGATAAGAAATGAAAAGGATGAAAAGGATGAAAAGCAACGCTGGGTGCGTATCACGCC 5220
Db 5161 GTATTGTAGATAAGAAATGAAAAGGATGAAAAGGATGAAAAGCAACGCTGGGTGCGTATCACGCC 5220
Qy 5221 CTGCTTCCAAAAGAAATTTTCTTTTAAAGTAAATCCCTAAATTTTCAATAACATCCGAGG 5280
Db 5221 CTGCTTCCAAAAGAAATTTTCTTTTAAAGTAAATCCCTAAATTTTCAATAACATCCGAGG 5280
Qy 5281 ATGCTGAGAACACCCGTCGCAAGTGTAGTGAATATCAAAAGGAGTGGCTATGGAAGAAG 5340
Db 5281 ATGCTGAGAACACCCGTCGCAAGTGTAGTGAATATCAAAAGGAGTGGCTATGGAAGAAG 5340
Qy 5341 GATACGTCTCTTATCTTTGGAGTTCGTTTCAATTTGTGTAGTACATAAAAATATATGTAA 5400
Db 5341 GATACGTCTCTTATCTTTGGAGTTCGTTTCAATTTGTGTAGTACATAAAAATATATGTAA 5400
Qy 5401 GAAAAGGTTTGAGGGAACGTAATTTTGAGTGTGACAGACGCTCGCCAAATTCGAACCTCCTG 5460
Db 5401 GAAAAGGTTTGAGGGAACGTAATTTTGAGTGTGACAGACGCTCGCCAAATTCGAACCTCCTG 5460
Qy 5461 AAAAGGTTTGAGGAGTTCGTCGATGAAAGTACCAATGGCTGTGAAAACCTCGAAAAGGTTTC 5520
Db 5461 AAAAGGTTTGAGGAGTTCGTCGATGAAAGTACCAATGGCTGTGAAAACCTCGAAAAGGTTTC 5520
Qy 5521 CGGAAAAACAAAAAGAAATGAGTAAATATATGTTAATAAAGAAATAAATAACAGTG 5580
Db 5521 CGGAAAAACAAAAAGAAATGAGTAAATATATGTTAATAAAGAAATAAATAACAGTG 5580
Qy 5581 GTAGAGAGGTTTAAATTTGAGGAAATTTAGGAGTAAATGTAAGTGTGACGAGTCTATCG 5640
Db 5581 GTAGAGAGGTTTAAATTTGAGGAAATTTAGGAGTAAATGTAAGTGTGACGAGTCTATCG 5640
Qy 5641 CGTCATCGAGTACGTTTTTAAATCAATATGCTTTTATCAATCAACTCTCCGAGCCAAATTTGT 5700
Db 5641 CGTCATCGAGTACGTTTTTAAATCAATATGCTTTTATCAATCAACTCTCCGAGCCAAATTTGT 5700

Qy 558 AAAGGAGGCTTTTAAACAGGTATGACAGAGCTCCCAACGAAGTCTGCTGCTCTTAAACTT 617
Db 559 AAAAGAGGCTTTTCGACAGATACGCTGAAATGCCAAAACGAAGTAGTCTGTCCACGATACTT 618
Qy 618 TTCAAGGATGTCGAATACATCCGCGACAGAGATAGTGTGAGAGATACGCTGTGCTCTGC 677
Db 619 TCCAAACGTGTAGCATCTCAAGAAATGTTACACGGAGAGTGTATGCTATTGCTTTGC 678
Qy 678 ACAGTTTGTATGATATTTCTGTGTCATGAGTTTGGAGTGGGTGATATATCTAAGAAATATAC 737
Db 679 ATAGTATATACGATATACCTGCGACGAGTTTCCGCGGCGCACTGCTCGAAGAAATGTAC 738
Qy 738 ATGATCTTATGACAGCTTCCATTTTGGCAGAGCAATTAATTAAGCAGACGAGGTAA 797
Db 739 ATGATGTATTGCGCGCTTTCCACTTTTCCGAGAAATTTACTTCTCGAAGATTCACACGTCA 798
Qy 798 CGCTTAATGAATAGGCGCACTTTCAAGAGAGAGGTGATGATGTTCTTTTCTTTTCTTTG 857
Db 799 ACCTCGACGAGATCAATGTCATGTTTCCAAAGAGATGGAGCAGGTGACITTTTCTTTG 858
Qy 858 CTGATGAAGTACTTTTAAATATATAGTCATAAATCAAAAAATATCTTGCATTTATGTAGTTA 917
Db 859 CATCTGAGAGTACTTCTAAATATATAGTCATAGTTATCTTAATTTCTTAAGTATGTTGCA 918
Qy 918 AATCTTACTTCTGCTCTAGTAGAATAGTTTACTTTAAGGAATTTTACTACTAGG 977
Db 919 AAACCTTACTTCCAGCCTCTAATAGAGAGGTTTACATGAAGGAGTTTTTATGTAATAGAG 978
Qy 978 TTAATACTGTGTTTGTAAATTTTACCAGATAGATACCTATATCTGTACAAAGAGTGTAA 1037
Db 979 TTAATACTGTGTTTGTAAATTTTCTAGATAGATATCTTCTTATTTAGTACAAAGGTGTAG 1038
Qy 1038 GACAAGTAGGTTGTATGATGATCAGTTCTATGAGGCGATGGAAGAGCCCTTTGCTTTACA 1037
Db 1039 CGCATAGAGGTGTAGATAGTACGAGTTTACAAAGGCTATGGAAGAGCAGTGGCACTACA 1098
Qy 1098 AGAAACCTTGGCCATGTTCAACACTGAAAGAGCAATCTTTAGAGACAGCGCTTCGGTTA 1157
Db 1099 AAAAGACTTCTGCGATGTGCAACAGTGAAGAAATCTTTGTTAGAGAGATTTCTTCATCAGTTA 1158
Qy 1158 ACTTTTGGTTCCTTAAGATGAAGACATATGAGTAGTACCGCTGTTGAGGGTTCTATTAA 1217
Db 1159 ATTACTGTTTCCAAAAATGAGGATATGTTGATAGTTCCTATTTTGACATATCTCTCG 1218
Qy 1218 CCAGCAAAAGATGACAAGGAGTGAAGTCAATTTGTTAAATCGTGACTTCGTTTACACAGTGC 1277
Db 1219 AGACTAGTAAAGAACACGCAAGAGGCTTAGTTTCAAAGGACTTTGTTTATACAGTGT 1278
Qy 1278 TTAATCATATCAGAACATATCAAGCCAAAGCGTTAACTTACAGAACGTTATTTCTTTTCG 1337
Db 1279 TAAATCATATTCGTACGTACAGGCCAAAGCGCTTACTTACTCCAACGTTATCTTTTCG 1338
Qy 1338 TGGAGTCTATAAGATCCGCGTGATATCAATGTTGTTACTGTAGGTCTCAATGGGATG 1397
Db 1339 TCGAATCAATTCGTTTCGAGAGTGATCAATACGGGTTACTGTAGTCTGAGTGGGATG 1398
Qy 1398 TAGATAAAGCAATTTCTCAACCTTGTCAATGACTTTCTTCTGAGACTAAGCTGGCTG 1457
Db 1399 TCGATAAATCATTTATACAGTCTTGTGATGACGTTCTTCTCTCACACCAAGCTTGCCG 1458
Qy 1458 CGCTTCAAGACGATATAGTAATGGAAAGTTTCGGTGTCTGGATGAAGACCACTTCTGNAC 1517
Db 1459 TTCTGAAAGACGATCTTTTGTATTAGCAAGTTTGCACTTGGACCAAAAATCTGCTCACAAAC 1518
Qy 1518 TTATTTGGATGAGTGGGCAAAATTTTGTGAAAGGTTTTTCCCACTATCAAGAGAGAT 1577
Db 1519 ATGTGTGGATGAGATTTCCCTAGCTTTTCGGCAATGCTTTCCCATCGATCAAGGAAGAT 1578
Qy 1578 TGGTGACAGGAAATTTCTCGATGTAGTGAAGATGCTCTGAAGATCAAGATCCAGATC 1637
Db 1579 TGATAAACCGGAACTGATCAAAATACGGAGAAATGCGTTAGAGATCAGGGTGCCCGATC 1638
Qy 1638 TGTATGTACATGGAAGACAGGTTCTGTAGTGAATACCAACGCTCTGAGGAGTTACCGC 1697

Db 1639 TTTAATGTCATTTTCCATGATAGGTTTAGTTTCTGAGTACAAAATGTCAGTGAACATGCCGG 1698
Qy 1698 ATCTAGATATCAAGAGAGACTTAGAAGAGCTGAGCAAAATGTTACGACGCGTTTATCAGAAAT 1757
Db 1699 TGCTAGACATTTAGGAAAAAGATGGAAGAACTGAGGAAATGTTACAAATGCACTGTCCGAAC 1758
Qy 1758 TATCTATCTTAAGGGTGTCTGATATTTTCGATATCGCGAAGTTTCAAGACATGTGCAAGG 1817
Db 1759 TGTCTGTACTTAAAAATTCAGACAAAGTTTCGATGTGACGTTTTTTCACAGATGTGCCAAT 1818
Qy 1818 CTTTATAGATTTAGTCTCGATGTGGCAGCACAGTAATTCGTTGACGTGAGTGGCCGAGAAATAGAA 1877
Db 1819 CTTTGAAGTCGATCCAAATGACTGACAGCAAAAGTTAATAGTACAGTTATGAGCAACGAGA 1878
Qy 1878 GCGGTTTAACTCTTACTTTTGAATAGCCAAACCGAGGAGAAATGTGGCTTAAAGCTCTTAA- 1936
Db 1879 GTGGTCTTACTCTCACGTTTGAACAGCCCAACCGAAGCTAATGTTGCGCTAGCTAGTCAAG 1938
Qy 1937 -----ACACGCGCTCTGAGCGCTGTGATGTTTGAACCGACATCCGAAAGAGTGAACG 1991
Db 1939 ATTCTGAAAAGGCTTCTGATGGGCGCTTGGTAGTTTACCTCAAGAGATGTTGAGGAACCGT 1998
Qy 1992 TAAATAAATTTTCTATTGCTGAGAAAGGAGATTTGCCCTGTGTGTCAGAAAAGTCAATGCTT 2051
Db 1999 CCATAAAGGTTTCATGCGCGCTGTGAGTTACAATTTGGCCGGATTTATCTGGCAGCTTC 2058
Qy 2052 TGACGAATGCTTAACCTTAGAGCACAGGAGTTGGAGTCCCTCAACGATTTTCCATTAAGGCTT 2111
Db 2059 CTGAATCTTCATACACTAGGAGCGAGGAGATTGAGTCTCTCGACAGTTTTCATATGGCAA 2118
Qy 2112 CGGTGAGTAGTGTGATTAACAAAGCAATGGCATCGGTGTGCTACACTGGCTCACTCAAG 2171
Db 2119 CAGCTAGTTCTGTTAAATTCATAAGCAGATGTTTCGATCGTGTACACGCGGCCCTCTTAAAG 2178
Qy 2172 TTCAACAAATGAAGAACTATGTGACAGATTTGGCAGCTTCGTTCTCGCCCACTGTATCAA 2231
Db 2179 TTCAACAAATGAAGAACTTTATAGACAGCTTGGTAGCTCGCTCTCTGCTCGGTGTGCA 2238
Qy 2232 ATCTATGCAAGTCTCAAAAGGATGAAGTCCGGTATGATTTCTGATTTCCAAGGAGAAAGTTG 2291
Db 2239 ATCTAGTGAAGATCTCTAAAAGATACAGCCGCGATTGACCTTTGAAAACCTCGTCAAAAGTTCCG 2298
Qy 2292 GTGTTTGGATGTCACTTTGAAAAGTGGCTCTCAAACTGCGGCCCAAGGTCATTTCTAT 2351
Db 2299 GAGTTCTGGATGTGCTTCGAAAAGGTGGCTAGTTAAACCACTCCGCAAGAACCATGTGAT 2358
Qy 2352 GGGGAGTTGCTCGATTACAAGGGGAAATGTTTACTGCACTTCTATCTTATGAAGGAG 2411
Db 2359 GGGGGTTTGTGACATCTGCGAGGAATATACAGTCCGATTTACTGGAGCACGATGAAT 2418
Qy 2412 ATAGAATGGTGACTGAGAGCGACTGGAGGAGGGTGGTGTATCATCTGATACAAATGGTAT 2471
Db 2419 TTGSCATTATTACGTGCGATAAATCTGGCGACGGGTGGCTGTGAGTTCTGAGTCCGTAGTAT 2478
Qy 2472 ATTCTGATATTTGCAAAAGCTCCAAATCTGAGGAAAAACAATGAGAGACGGTGAACCCACG 2531
Db 2479 ATTCTGATATGGCTTAAACTCAGGACTCTGAGAAAGTTGCTCAAAAGATGGAGAACCCACG 2538
Qy 2532 AACCTACTGCAAAAGATGTTACTGATGGGTGGCTGCTGTTGTTGAAAGATACAAAGGAG 2591
Db 2539 TTAGTTCAGCAAGGTGGTTTGTGATGGCTTCAGGGTGGGGAGAGACAAAGGAAA 2598
Qy 2592 ATTTTGAAGATTTGATCTTGATGAGGATTTGATCTTTGGTTTCTTGGAAAAACAAGTGTG 2651
Db 2599 TTTCTTCGAGAGTTAAATTTTCGAAGAAGATCTAAATTTCTTGTCCCTGGTCTCAAGCTGCCG 2658
Qy 2652 CTATGATCAGAAAGGGCTTAATTTCACTGCACTGATGAAGCCACATGACATGCAATGTGA 2711
Db 2659 AGATGATCAGAAAGAGCTAAATGCGTGGGCAATAATAGTGGCTACAAAGGATTAATGTGC 2718
Qy 2712 GAACGGTAGATTCTTCTTATGATCATCCAAAACCGCGATCAC-----ACAAGAGGC 2762

Db 2719 GCACCGTCGATCAATTTTGTGATGAATTAACGGAAAGGGGACACGCTGTCTCAAGTTCAAAAGAT 2778
Qy 2763 TTTTATTTGATGAAGGGTTGATGCTGCACACCGTTGTGTAACTTTCCTGGTCTTATCT 2822
Db 2779 TGTTCATAGACGAAGTTTGTGCTGCATCTGTTGTGTGTAATTTTGGTTGAAATGT 2838
Qy 2823 CTGGTTGCGACATCGCATACATTTACGGAGATACACAGCAGATTCCTTTTCATTAACAGAG 2882
Db 2839 CTCGTGCGATATGCAATATGTTATGGAGACACCCAAACAGATTCGTACATCAACAGAG 2898
Qy 2883 TTCAGAAATTTCCGGTATCCAAACATTTTGAAGCTGCAAGTGAGTGAAGTGAATGA 2942
Db 2899 TAACTGGTTTCCCGTACCCCTGCACACTTTGCAAAATTTGAGGTGACGAAAGTCGAAACAA 2958
Qy 2943 GGAGGACCACTGAGATCCCGAGGTGATGCAATTTTTCCTACAATCGAAGTACGAG 3002
Db 2959 GAAGAACTACTCTTGGTGTCCGGTGTATGCACACTTCCTTAATCAAGGTATGAAG 3018
Qy 3003 GAGCGGTGACAACTCTCAACTGTAACAACGATCGGTCTCATCTGAGATGATPAGCGGTA 3062
Db 3019 GACACGTAATGTGCACGCTCTCTGAAAGAAATCAGTTTCCAGGAAATGGTTAGTGGG 3078
Qy 3063 AGGAGTACTAAACAGTGTTCACCAACCACTAAAGGGGAAATTTGTAACCTTCACTCAGG 3122
Db 3079 CTGCGTCTATCAATCCTGTGTCCAAAGCGCTTAAAGGGGAAATTTTGAATTTTCAACAGT 3138
Qy 3123 CTGATAAATTTGAGTTAGAGGAGAGGGCTATAAGAAATGTGAACACCGTTCAATGAGATCC 3182
Db 3139 CTGACAGGAGGCCCTCTCTCAAGGGCTAGCGAGATGTCATACTGTACATGAGGTAC 3198
Qy 3183 AAGGAGAACTTTGAAGATGTGCGGTGAGATTTGACGGCAACTCCACTGACTCTGA 3242
Db 3199 AAGGTGAGACTTATGCAGACGCTATCGTTAGTTCGACTAAACCTACGCTGTATCTATCA 3258
Qy 3243 TTTTCAAGTCTTCCCGCATGTTCTAGTCGCTCTGACTAGACACACAAGAGCTCAAT 3302
Db 3259 TCGCAAGAGACAGTCCGATGTTCTGGTCTCGTTGTCAAGACACACAAAATCCCTAAAGT 3318
Qy 3303 ATTACACCGTAGTTAGATCCTTTAGTACAGATAATTAGTGTGTTCTCTTTAAGCT 3362
Db 3319 ACTACACCGTTGTGATGATCCTTTAGTTAGTATCATTAGAGATTTAGAACGGGTAGTA 3378
Qy 3363 CTTCTCTTTTGAATATATGTTAGTAGAAGCGAGTAGTAGATAGCAATTAACAGATGAGT 3422
Db 3379 GTTACTTATTAGACATGTACAAGTAGATGACAGGTACTCAATAGCAATTAACAGGTGACT 3438
Qy 3423 CAGTGTTCAAAGGTCAATCTTTTGTGGCAACACTAAATCAGGAGACTTCCAGATC 3482
Db 3439 CTGTTTAAAAATTTCAATCTTTTGTAGCAGCTCCAAAGACTGGAGATATATCTGATA 3498
Qy 3483 TACAGTCTTATPACGATGTATGCTCCTCGTGAATAGTACTATATTAACCAAGTATGATG 3542
Db 3499 TGCAATTTTACTATGATAAGTGTCTCTCGGAACAGCACGTTGTTGAACAACTACGACG 3558
Qy 3543 CTGTTACCATGAGGTTAGTGATAATAGTCTTAATGTGAAGATGTGTTCTGATTTTT 3602
Db 3559 CTGTTACCATGAAATTTGACTGACATTTCTCTGAATGTCAAAGATTTGCAATATAGATGT 3618
Qy 3603 CCAAAAGTATTCGGATGCAAGGAGTGAAACCAATGTCTAGACCACTTTTGGTACCG 3662
Db 3619 CTAAGTCTGTAGTCTCTCGAAGATGTCAAACCACTTTAATCCGATGTGCAACGG 3678
Qy 3663 CGCGGAACCGCGAGCTGCAGGACTTACGAAATCTGGTTGCAATGATTTAAAGAA 3722
Db 3679 CGGCAGAAATGCTCGCCAGACTGACCTGTTGGAAATCTAGTTGCGATGATTTAAAGAA 3738
Qy 3723 ATTTCAACGACCACTGAGGAGTGTGATGATGATGATGATGATGATGATGATGATGATG 3782
Db 3739 ATTTTAAATCCACGAGTGTGCGGAGTGTGATGATGATGATGATGATGATGATGATGATG 3798
Qy 3783 TAGATAAGTTTTTGTAGTATTTTATTTAAAGAAAGAAATATACAAAAAATATTTGCTG 3842
Db 3799 TAGATAAGTTTTTGTAGTATTTTATTTAGGAAAGAAAGAAAGAAAGAAAGAAATTTT 3858

Qy 3843 GAGTATCACGAAGATTTCAATGATGAGATGCTTGGAAAAACAGGAAAGAAAGTACTATTGG 3902
Db 3859 CACTGTTTAGTAGAGTCTCTCAATAGTGTGATAGCAAGCAAGCAAGTCAAAATG 3918
Qy 3903 ACGACTTGGCTAACTACAAATTTTACAGATCTGCGCGCCATCGATCAGTACAGACACATGA 3962
Db 3919 GTCAGTTGGCCGATTTTGTGATCTTCAGCCGTTGATCAGTACAGGCATATGA 3978
Qy 3963 TCAGGCTCAACCAAAACAGAAATTTGGACCTTTCAATTCAGAAATGAATACCCCTGCTGTC 4022
Db 3979 TTAAGCCCAACCGAAGCAGAACTGGATCTGTCAATTCAGACAGAAATATCCAGCGTTGC 4038
Qy 4023 AAACAATTTGCTACCATTCGAAGCAGATCAACGGTATTTTGG--CCGGTTTCTCAGAG 4079
Db 4039 AAACGATTTGTATCAITCAAAGAAATCAAACGCAATATTTGGTCTCTTTTCAGTGAGC 4098
Qy 4080 TTACAAGTTTCTGCTCGAGGCAATTTGATTTCTAAGAAGTTTCTTTTCTTACTAGGAAAA 4139
Db 4099 TTACAAGCAATTTACTTGACAGTATTGACTCAAGCAGATTTCTTCTTTTACGAGAAAGA 4158
Qy 4140 CTCCAGAACAGATTCAGAAATTTTCTCGGATCTGACTCGACGTTCCCTATGATGTGT 4199
Db 4159 CACCGGCTCAGATCGAAGATTTCTTCGGAGATCTAGACAGTCATGTCCCAATGACGCTAC 4218
Qy 4200 TAGAACTGGATATTTCTAAGTATGATAGTCACAGAACGAGTTTTCATTTGCTGTAGAGT 4259
Db 4219 TTGAGTTGGAATTTTCCGAGTATGATAGTCTCAAAACGAGTTTTCATTTGCTGTGAGT 4278
Qy 4260 ATGAAATATGAAAGATTTGGGTCTCAATAGTATTTTGGCCGAAAGTGTGGAAAAACAAGGC 4319
Db 4279 ACGAAATCTGGAGGAGACTGGGTCTGGAGGATTTCTTGGCAGAAAGTGTGGAAAAACAAGGC 4338
Qy 4320 ACAGGAAACAACTTTCGAAGATTTACATTCGAGAAATCAAGACATGCTGTGATCAAA 4379
Db 4339 ATAGAAACCACTCTGAAAGATTTACATCTGTGTATAAAACCGTGTATTTGATGCCAGA 4398
Qy 4380 GGAAGGCGGTGATGTGACTTCTTTCATCGCAATACTGTTATATATAGCAGCTGCTTGG 4439
Db 4399 GAAAGAGTGTGTATGTTACAACTTTTATCGGTATACCGTCATCATTTGCTTCTGCTAG 4458
Qy 4440 GTTCAATTTACCAGTGAAGGATCAATAAGGTGCTTTTGTGGAGACGATTTCCGTTT 4499
Db 4459 CATCAATGCTCCCGATGGAATAATGATAAAGAGAGCTTCTCGCGGAGATGACAGTTTGT 4518
Qy 4500 TGTATTTCCAAAGGGTTTGGATTTTCCCTGACATTCAGTCAATGCTTAATCTCATGTGA 4559
Db 4519 TGTACTTTCTAAGGGTTGTAGTATCCCGATATACAAACAGCTGCTTAATCTAATGTGGA 4578
Qy 4560 ATTTTGAAGGCCAACTGTATAGAAAGAGGTACGGTTACTTTTGTGGTAGATACATCATAC 4619
Db 4579 ATTTTGAAGGCCAACTGTTCAGAAAGCAATATGSGTACTTCTCGCGGAGGTACGTGATTC 4638
Qy 4620 ACCATGATAGAGGAGCAATAGTGTATATGATCTCTTTGAAAGTGAATCTCCAAACTGGGG 4679
Db 4639 ATCAAGATAGAGGTTGATAGTATCTACGACCTTTTGAAGCTGATTTCCGAACTTGGTG 4698
Qy 4680 CAAACATATCAAGGATTTATGATCCTACATAGAGGTTTAAAGGTGCTTTTGTGCGATGTTG 4739
Db 4699 CTAAACATATCAAGATTTGGATCATTTGGAGGAGTTCAAGATCCCTCTGTGATGTTG 4758
Qy 4740 CTTGTTCCGCTCGGAACTGGTGTCTTAGGCTTTCCGACAGCTGAACGACGATATCAAGGAG 4799
Db 4759 CTGAGTCTGT--GAACAAATTCGCGTATTACACAAATTTGGACGACGCTGTTGGGAGG 4815
Qy 4800 TTCAATAAACCGGATTTGATGCTGTTGCTTTTAAATGTTGTTAAACAAATTTTGTGTG 4859
Db 4816 TTCAATAAACCGCCCCACCTGGTTCGTTGTTTATAAGAGTTTAGTTAAGTATTTTGTGAG 4875
Qy 4860 ATAAATTTTATTTAGAACTTTGTTTTTAAATGGCTGTAGTCTCAGAGATACGTGCAAA 4919
Db 4876 ATAAAGTTTTTGTATTAGAGTTTATTTCTTGTGCTCTAGTTGTTTAAAGGTAGGTTAAT 4935


```
QY 4920 ATTACGAGTTCAATGATCTTTGAAACAGGATGAGATACCTCCGGCATTCATGACTAAG 4979
DB 4936 ATTAATGAGTTATCGATCTGTCAAGTCTGAGAAACTTCTCCGCTGATGTTCAACGCT 4995
QY 4980 GTCAAGAGTGTAGAAATACCACTGTGACAAAGATTATGGCTGTGTAAGAAATGATAGTCTT 5039
DB 4996 GTAAAGAGTGTATGGTTTCAAGGTTGATAGATTATGGTCCATGAAATGAATCATTTG 5055
QY 5040 TCTGATGAGATTTACTTAAAGGTGTTAAAGTTAGTTAAGTAAAGAAAGGTATGCTGTCTTAGCT 5099
DB 5056 TCTGAAGTAAATCTCTTAAAGGTGTAATACTTATAGAAAGGTGGGTATGTTGCTTAGTC 5115
QY 5100 GATTGGTAGTGTCTGGGAGTGGATCTCCGATTAATCTCCGCTGAGTGTGAGTGT 5159
DB 5116 GGTCTGTGTGTCCGGTGGTGGAAATTTACAGATTAATCCGCTGGTGGTGTGAGTGT 5175
QY 5160 TGTATTGTAGATAGAGAAATGAAAGAGTAGTAAGGAAGCAAGCTGGGTGGCTATCACGCC 5219
DB 5176 TGCATGTTGACAGAGAAATGGAAGAGCGGACGAGCCACACTGGGTGATATTTACACT 5235
QY 5220 CTTGCTGCAAAAAAGAAATTTTCTTTTAAAGCTAATCCCTAAATTAATCAATAACATCCGAG 5279
DB 5236 CTGCTGCTGCAAAAAAGCGTTTCAGTTTAAAGTGGTCCCAATTTACGGTATTTACAACAAAG 5295
QY 5280 GATGCTGAGAAAGCAACCGGTGCAAGTGTAGTGAATATCAAGAGAGTGGCTATGGAAGAA 5339
DB 5296 GATGCAAGAAAGAACATATGCGAGCTCTTAGTAATATTAATAATTAATAAATGTAATAAGTGTGCG 5355
QY 5340 GGATACTGCTCTTATCTTTGAGTTCGTTTCAATTTGTTAGTAGTACATAAAAAATAATGTA 5399
DB 5356 GGCTACTGCCCTTTGCTCAATTAGAAATTTGTTGCTGTGTGTTATGTTTATAAAAAATATA 5415
QY 5400 AGAAAGGTTTGGAGGAAACGTATTTTGTAGTGTGACAGACGGCTCGCCAAATTTGAACCTACT 5459
DB 5416 AAATTTGGGTTTGGGGAAGAAATGTAACGAGTGTGAACGATGAGGACCCATGAACTTTTCG 5475
QY 5460 GAAAAGTTTGTAGGAGTTTCGTGATGAAGTACCAATGGCTGTGAAACTCGAAAAGTTT 5519
DB 5476 GAAGAAGTTTGTAGTGTTCATGAGAAATGTTCCAAATGTCGGTTAGACTCGCAAGTTT 5535
QY 5520 CCGGAAACAAAAAGAAATGTAAGTAAATGTTTAATAT----- 5561
DB 5536 CGAACCAATCTCMAAAGAGTCCGAAAATAATAATAATTTAGTTAAGGGGGGTTCAC 5595
QY 5562 ---AAGAAAATAATAACAGTGTGAAGAGGTTTAAATAATTGAGGAAATTTGAGGATAAT 5618
DB 5596 GGCGGAAGGCTTAAACCAAAAGTTTGTATGAAGTTGAAAAGAGTTTGTATAATTTGATT 5655
QY 5619 GTAAGTGTAGAGTCTATCGGTCATCGAGTAGCTTTTAATCAATATGCTTATACAA 5678
DB 5656 GAAGATGAAGCGGAGACGTCGGTCGCGGATTTCTGATTCGTATTAATATGCTTACTCAA 5715
QY 5679 TCAACTCTCCGAGCAATTTCTTTTACTTATCTTCGGCTTACGAGATCCTGTGAGCTGA 5738
DB 5716 TCACCTCTCCATCGCAATTTGTGTTTTGTATCTGTATGGGCTGACCTATAGAAATGT 5775
QY 5739 TCAATCTGTGTACAAATGTCATTTGGGTACCACTGTTTCAACGCAACAAAGCTAGGCAACAG 5798
DB 5776 TAAACGTTGTACAAATTCGTTAGTAACTAGTTTCAACACAGCAAGCAAGCACTACTG 5835
QY 5799 TCCAAACAGCAATTTGCGGATCGCTGGAAACCTGTGCTAGTATGATGAGTGAATTTCTGT 5858
DB 5836 TTCAACAGCAGTTTCAAGCAGGTGTGGAAACCTTTCCCTCAGACGACCGCTCAGATTTCTGT 5895
QY 5859 CATCGGATTTCTATGTGTATAGATATAATTCGACGCTTGTATCGCTGTATGATCAGGGTTAT 5918
DB 5896 GCGATGTTTAAAGGTGTACAGGTACATGAGTTTGTAGATCCTCTAATTTACTGCGTTGC 5955
QY 5919 TAAATAGCTTTGATACTAGAAATAGATAATATAGAGTTTGATAATCAACCCGACCGCAATA 5978
DB 5956 TGGGGCTTTTGTACTAGTAAGTAAGTAATTCGAAGTAGAATAACAGCAGAGTCCGACAA 6015
QY 5979 CTACTGAATCGTTAAACCGGACTCAGAGGGTAGACGATGCTACTGTACTATAAGGGCTT 6038
```

```
DB 6016 CAGCTGAAACGTTAGATGCTACCCGACGGGTAGACGAGCTACGTTGCAATTCGGTCTG 6075
QY 6039 CAATCAATAAATTTGGCTAAATGAATGGTTTCGTGGAACCTGGCATGTTTCAATCAAGCAGGCT 6098
DB 6076 CTATAAATAAATTTAGTTAAATGAATAGTAAAGGTTACTGGACTGTACAAATCAGATACTT 6135
QY 6099 TTGAGACTGCTAGTGGACTTGTCTGGACCAACAATCCGGCTACTTAGCTATTGTTGTGAG 6158
DB 6136 TTGAAGTATGCTCGGGTTGGTCTGGACCTCTGCACCTGCATCTTAAATGATAGTGTCT 6195
QY 6159 ATTTCTTAAATAAAGTCGCTGAAGACTTAAATTCAGGGTGGCTGATACCAAAATCAGC 6218
DB 6196 GAAATATAAA-----GTTTGTGTTTCTAAAAACACACGTTGACGTACGATAACGTAC 6247
QY 6219 AGTGGTTGTTTCGTCACCTTAAATATAACGATTTGTTCATATCTGGATCCAAACAGTTAAACCA 6278
DB 6248 AGTGTGTTTTCCTCCACTTAAATCGAAGGTAGTGTCTTTGGAGCGCGGAGTAACATA 6307
QY 6279 TGTGATGGTGTATATCTGTGTATGCGGTAAACATCGGAGAGGTTTGAATCCTCCCTTAA 6338
DB 6308 TATGGTTTCATATATGTCGTAGGCACGTAAAAAAGGAGGATTCGAATTTCCCCCGGAA 6367
QY 6339 CCGCGGTTAGCGGCCA 6355
DB 6368 CCCCCTGGTGGGCCCA 6384

RESULT 4
ABA04211
ID ABA04211 standard; DNA; 6384 BP.
XX
AC ABA04211;
XX
DT 01-MAR-2002 (first entry)
XX
DE Tomato mosaic virus related DNA sequence.
XX
KW Tomato mosaic virus; vaccine K; low-virulent vaccine K genome;
KW point mutation; replicase; motor protein; plant virus genome;
KW genetic engineering; viral disease; plant; ds.
XX
OS Tomato mosaic virus.
XX
PN CN1306090-A.
XX
PD 01-AUG-2001.
XX
PF 14-JAN-2000; 2000CN-00100211.
XX
PR 14-JAN-2000; 2000CN-00100211.
XX
PA (MICR-) INST MICROBIOLOGY CHINESE ACAD SCI.
PI Qiu B, Yang G, Tian B;
XX
WPI; 2002-000227/01.
XX
PT Low-virulent vaccine K genom sequence of tomato mosaic.
PS
CC Example 4; Fig 1 (Disclosure); 22pp; Chinese.
CC
CC The present invention describes the genome separated and cloned from
CC tomato mosaic virus, which is modified into low-virulent vaccine K genome
CC by point mutation. The nucleotide sequence of its coded replicase and
CC motor protein has the mutation, so it can be used to modify the plant
CC virus genome and prepare genetically engineered vaccine for effectively
CC preventing and treating the viral diseases of plants. The present
CC sequence represents a DNA sequence which is given in an example from the
CC present invention
XX
SQ Sequence 6384 BP; 1906 A; 1185 C; 1478 G; 1815 T; 0 U; 0 Other;
```

[illegible]

Db 2119 CAGCTAGTTCCTTAATTCATTAAGCGAGATGTGTTTCGATCGGTACACGGGGCCCTCTTAAAG 2178
Qy 2172 TTCAACAAATGAAGAACTATGTGGACAGTCTTGGCGAGCTTCGTTGCGCCACTCTGTATCAA 2231
Db 2179 TTCAACAAATGAAGAACTTTATAGACAGCTGGTAGCCTCGCTCTGCTGCGGTGCGA 2238
Qy 2232 ATCTATGCAAGTCTACTAAGATGAGTTCGGGTATGATCTGATTTCCAGGAGAAAGTTG 2291
Db 2239 ATCTAGTGAAGATCCTTAAAGATACAGCCGCGATTGACCTTGAACCTCGTCANAAAGTTG 2298
Qy 2292 GTGTTTGGGATGTCACTTTGAAAAAGTGGCTCCTCAAAACCTGCGCCCAAGAGTCAATTCAT 2351
Db 2299 GAGTTCTGGATGTGCTTCGAAAAGTGGCTAGTTAAACCAATCCGCAAGAGCAATGCAAT 2358
Qy 2352 GGGGAGTGTCTCTGATTAACAAGGGGAAAAATGTTTATCTGCACTTCTATCTTATGAAGGAG 2411
Db 2359 GGGGGTGTGTTGAGACTCATCGAGGAAATATCGCGTCGCATTAATCTGGAGCACGATGAAT 2418
Qy 2412 ATAGAAATGGTCACTGAGAGCACTGGAGGAGGTGGCTGTATCATCTCATACAAATGGTAT 2471
Db 2419 TTGGCAATTAATACGTCGCAATAACTGGCGACGGGTGGCTGTGAGTTCTGAGTCCGTTAGTAT 2478
Qy 2472 ATTCTGATTAATGCAAAAGCTCCAAATCTGAGGAAAAAACAATGAGAGCGGTGAACCCACG 2531
Db 2479 ATTCTGATATGCTAAACTCAGGACTCTGAGAAGATTGCTCAAGATGGAGAACACACG 2538
Qy 2532 AACCTACTGCAAGAATGCTACTTGTGGATGGGGTGCCTGTTGTGGAAAGTACAAAGGAG 2591
Db 2539 TTAGTTCAAGCAAGTGGTTTTGTGGATGGCTTCCAGGGTGGCGGAAGCAAGGAAA 2598
Qy 2592 ATTTTGAAGATTGATCTTGATGAGGATTGATCTTGGTTCTGGAAAAACAAGTGTCTG 2651
Db 2599 TTCTTTCAGAGATTAAATTCGAAGAAGACCTAAATCTTGTGCTCGTCAAGTGTCCG 2658
Qy 2652 CTATGATCAGAGAGGCTAATTCATCTGACTGATTAAGAGCCCAATGACATGTA 2711
Db 2659 AGATGATCAGATGAAGACTAATGGTCGGGCATTAATAGTGGCTACAAGGATTAATGTC 2718
Qy 2712 GAAAGGTAGATTCACTTCTAATGCAATCCAAAACCGCGATCAC- - - - -ACAAGAGC 2762
Db 2719 GCACCGTCGATTCATTTTTTGATGAATACGGGAAAAAGGCGCGCTGCAGTTCAAAAGAT 2778
Qy 2763 TTTTATTTGATGAAGGTTGATGTGCAACCGGTTGTGTTAACTTCTCGTGTCTATCT 2822
Db 2779 TGTTTCATAGACGAAGTTTGTGTCGATCTGCTGTTGTGTGAATTTTTTGTGTAATGT 2838
Qy 2823 CTGGTTGGACATCGCATACATTTACGGAGATACACACAGATTCCTTTCAATTAACAGAG 2882
Db 2839 CTCTGTGCGATTTGCAATATGTTTATGAGACACCCAAACAGATTCGCTACATCAACAG 2898
Qy 2883 TTCAGAAATTTCCGCTATCCCAACATTTTGAAGAGCTGCAAGTGGATGAAGTTGAGATGA 2942
Db 2899 TAACTGGTTTTCCGTTACCTGCACACTTTGCAAAATTTGGAGTGCAGCAAGTCGAACAA 2958
Qy 2943 GGAGAACCACTGAGATGCCAGGTGATGGAATTTTTTTCGTAATTCGAAGTACGAAG 3002
Db 2959 GAAGAACTACTCTTCGTTGTCCGCTGATGTACACACTTCTTAATCAAAAGGTATGAAG 3018
Qy 3003 GAGCGGTGCACAACCACTTCAACTGTACAAACGATCGTCTCATCTCAGATGATAGCGGTA 3062
Db 3019 GACAGTAATGTGCACGCTTCTGAAAAGAAATCAGTTTCCAGGAATTTGTTAGTGGGG 3078
Qy 3063 AGGGAGTACTAAACAGTGTTCCAAACCACTAAAAAGGAAAAATTTGAATTTTCACTCAGG 3122
Db 3079 CTGGCTCTATCAATCTGTGTCAGCGCTTAAGGAAAAATTTTGAATTTTCAACAGT 3138
Qy 3123 CTGATAAATTTGAGTTAGGAGAGAGGCGTATAAGAAATGTGAACACCGTTCAATGAGATCC 3182
Db 3139 CTGACAGAGGCGCCCTTCTCAGAGGGCTACGCAATGTCCATCTGTACATGAGGTAC 3198
Qy 3183 AAGGAGAACCTTTGAAGATGTGCTGCTCAGATTCAGCGCACTCCACTGACTCTGA 3242
Db 3199 AAGGTGAGACTTATGACAGCGTATCGTTAGTTTCGACTAACCTTACGCTGTATCTATCA 3258

Qy 3243 TTTCCAAAGTCTTCCCCGCAATGTTCTAGTCGCTCTGACTAGACACACAAAAGAGCTTCAAAT 3302
Db 3259 TCGCAAGAGACAGTCCGCAATGTTCTGCTCTGTTGTCGAAGACACACAAAATCCCTAAAGT 3318
Qy 3303 ATTACACCGTAGTCTTGTAGATCCCTTTAGTACAGATAAATTTAGTGAATTTGCTTCTTTAAAGCT 3362
Db 3319 ACTACACCGTTGTGATGATCCCTTTAGTTAGTATCAATTTAGAGATTTAGAACGGGTTAGTA 3378
Qy 3363 CCTTCTCTTTTAGAAATGTATATGTTAGAGCAGGTAGTAGATAGCAATTTACAGATGGATG 3422
Db 3379 GTTACTTTATTAGACATGTACAAAGTAGATGACAGGTACTCAATAGCAATTTACAGGTCGACT 3438
Qy 3423 CAGTGTCAAAGGTCATAATCTCTTTTGTGGCAACACCTAAATCAGGAGACTTTCCAGATC 3482
Db 3439 CTGTGTTTAAAAATTTCAATCTTTTGTAGCAGCTCCAAGACTGGAGATATATCTGATA 3498
Qy 3483 TACAGTTCTATTAACGATGTATGCTCCCTGCTGAATAGTACTATATCTTAAACAGATATGATG 3542
Db 3499 TGCATTTTACTATGATAAGTGTCTTCTGGGAAACAGCACGTTGTTGAACAACTACGACG 3558
Qy 3543 CTGTTACCATGAGGTTTACGTGATAATAGTCTTAATGTGAAGGATTTGTTCTTGTATTTT 3602
Db 3559 CTGTTACCATGAATTTGACTGACATTTCTCTGAATGTCAAAGATTTGCATATTTAGATATGT 3618
Qy 3603 CCAAAAGTATTTCCGATGCGCAAGGAGGTGAACCAATGTCTAGAGCCAGTTTTCGCTACCG 3662
Db 3619 CTAAGTCTGTAGCTGCTCCGAAAGATGTCAAACCAACTTTAATACCGATGTTGTCGAACGG 3678
Qy 3663 CGCGGACCGCCAAAGGCTGACAGACTACTCGAAATCTGGTTGCAATCTGGTTGCAATGATTAAGA 3722
Db 3679 CGGAGAAATGCCCTCGCCAGACTGGACTGTTGGAAAAATCTAGTTGCGGATGATTAAGA 3738
Qy 3723 ATTTCAACGACACAGACCTGACGGGGAGTTGACATTTGAGAGCAGCCGATCTGTTGTAG 3782
Db 3739 ATTTTAAATTCACAGAGTTGTCGAGTAGTTGATATTGAANAATCTGCACTCTTTAGTGG 3798
Qy 3783 TAGATAAGTTTTTGTAGTATTTTATTTAAAGAAAAAATACACAAAAAATATTGCTG 3842
Db 3799 TAGATAAGTTTTTGTAGTATTTTACTTTAAGGAAAAAAGAAAAACCAACAAAAATTTT 3858
Qy 3843 GAGTGATGACGAAGGATTCATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3902
Db 3859 CACTGTTTAGTAGAGAGTCTCTCAATAGTGGATAGCAAGCAAGCAAGTACCAATG 3918
Qy 3903 ACGACTTGGCTAACTACAAATTTACAGATCTGCGCGCATCGATCAGTACAGCAATGA 3962
Db 3919 GTCAGATGGCGATTTTGATTTTGTGGATCTTCCAGCGTTGATCAGTACAGGCATATGA 3978
Qy 3963 TCAAGGCTCAACCAAAACAGAAATTTGACCTTTCAATTCAGAAATGAATACCTGCTCTGC 4022
Db 3979 TTAAAGCGCAACCGAAGCAGAACTGGATCTGTCAATTCAGACAGAAATATCCAGCGTTGC 4038
Qy 4023 AAACAATTTGCTCACTCCATTTGCAAGCAGATCAACGGTATTTTGG- - -CCGTTTCTCAGAGC 4079
Db 4039 AAACGATTTGTATCATTTCAAGAAATCAAGCAATATTTGGTCTCTTCTTTCAGTGAGC 4098
Qy 4080 TTACAAGGTTCTGCTCGAGGCAATTTGATTTCTAAGAAGTTTCTTTTCTTACTAGGAAA 4139
Db 4099 TTACAAGGCAATTTACTTGACAGATTTGACTCAAGCAGATTTCTTGTCTTTACGAGAA 4158
Qy 4140 CTCAGAACAGATTCAGAAATTTTTTCTCGGATCTCGATCTCGACGCTTCCTATGGAATGT 4199
Db 4159 CACCGCTCAGATCGAAGATTTCTACGGAGATCTAGACAAATCATCTCCCAATGGACGTAC 4218
Qy 4200 TAGAACTGGATTTTCTAAGTATGATAGTACAGACAGTTTTCATTTGCTGCTGAGT 4259
Db 4219 TTGAGTTGATGTTTTCGAAGTATGATAAGTCTCAAAACAGTTTTCATTTGCTGTTGAGT 4278
Qy 4260 ATGAAATATGAAAAGATTTGGGTCTCAATGAGTTTTTGGCCGAGTGTGGAAAAAAGGCG 4319
Db 4279 ACGAAATCTGAGGAGACTGGGTCTGGAGGATTTCTTGGCAGAGTGTGGAAAAAAGGCG 4338

QY 4320 ACAGGAAACAACTTTGAAGGATTAACATTTCTGCGAATCAAGACATGTCGTGGTATCAAA 4379
DB 4339 ATAGAAAAACCACTCTGAAGATTAACATCTGGTATATAAAACGCTTTTATGGTACCAGA 4398
QY 4380 GGAAGAAGCGGTGATGTGACTTTTCATCGCAATACGTTATATAGCAGCTTCTTGG 4439
DB 4399 GAAAGAGTGGTGAATGTTACACTTTTATCGGTAATACCGTCATCATTTGCTCGTCTAG 4458
QY 4440 GTTCAATGTTACCGATGGAAGGTCATAAAAAGGTCTTTTGTGGAGACGATTCGGTTT 4499
DB 4459 CATCAATGCTCCCGATGGAATAATGATAAAGGAGCCTTCTCGGAGATGACAGTTTGT 4518
QY 4500 TGTATTTTCCAAAGGTTTGGATTTCCCTGA CATTGACGTATGTCATACTCATGTGA 4559
DB 4519 TGTACTTTCTTAAGGTTGTGAGTATCCCGATATACAAAGCTGCTAATCTAATGTGA 4578
QY 4560 ATTTTGAAGCCAACTGTATACAAAGAGGTAGCGTTACTTTTGTGGTAGATACATCATAC 4619
DB 4579 ATTTTGAAGCCAACTGTTCAAGAGCAATATGGGTACTTCTCGCGGAGGTACGTGATTC 4638
QY 4620 ACCATGATAAGGAGCAATAGTGTATATGATCCTTTGAAAGTTGATCTCCAAACTTTGGG 4679
DB 4639 ATCAGATAGAGTTGCTAGTATATACTACGACCCCTTGAAGCTGATTTGAAACTTTGGT 4698
QY 4680 CAAAACATATCAAGGATTAATGATCACTTAGAAGAGTTAAGGTTGCTTTGCGGATGTTG 4739
DB 4699 CTAAACACATCAAGGATTTGGGATCATTTTGGAGGATTCAGAAAGATCCCTCTGTGATGTT 4758
QY 4740 CTGTGTTCCGTCGAAACTGGTCTTAGGCTTTCCGACGTGAACACGCTATCAAGGAGG 4799
DB 4759 CTGAGTCGTT---GAACTATTCGGGTATTACACAAATTTGACACGCTGTGGGGAGG 4815
QY 4800 TTCTATAAAACCGCGATTTGATGTTGCTTTGCTTTTAAATTTGTTTAAACAAATTTTGTGTG 4859
DB 4816 TTCTATAAAACCGCCACCTGTTGCTTTGTTTATAAGAGTTTGTAAATTTTGTCTAG 4875
QY 4860 ATAAATTTTATTTAGAACTTTGTTTAAATGGCTTTGTTCTCAGAGATCTGTCAAA 4919
DB 4876 ATAAAGTTTGTTTAGAAATTTATTTCTTGATGGCTCTAGTTTGTAAAGGTAAGGTAAT 4935
QY 4920 ATTAGCGAGTTCATTTGATCTTTTCGAAACAGGATGAGATCTTCCGCGATTTCATGACTAAG 4979
DB 4936 ATTAATGAGTTTATCGAICTGTCGAAAGCTGAGAACTTCTCCCGTCGATGTTTACGGCT 4995
QY 4980 GTCAAGAGTGTTPAGAATATCGACTGTGGAACAGATTTATGGCTTTTAAAGATGATAGTCTT 5039
DB 4996 GTAAAGAGTGTATGTTTCAAGGTTGATAAGATTATGGTCCATGAAAAATGAATCATTTG 5055
QY 5040 TCTGATGTAGATTTACTTTAAAGTGTTAAGTTAGTTAAGAAAGGTATGTGTGCTTAGCT 5099
DB 5056 TCTGAAGTAAATCTCTTAAAGGTGTAAAACTTATAGAAGGTGGGTATGTTTGTCTTAGTC 5115
QY 5100 GATTTGGTGTGTCCTGGGAGTGAATCTCCCGGATAACTCGCGGTGGTGTCTAGTGT 5159
DB 5116 GGTCTTGTGTGTCGGTGTAGTGAATTTACAGATAATATGCGGTGGTGGTGTAGTGTCT 5175
QY 5160 TGTATTTAGATPAAGAGATTAAGAAAGGAGTAAAGGAAGCAACGCTGGGTGCGGTATCACGCC 5219
DB 5176 TGCATGTTTGA CAAGAGAATGAAAGAGCGGAGCAAGCCACACTGGGTCTATATTACACT 5235
QY 5220 CTGCTGCGAAAAAGAAATTTTCTTTTAAAGCTAATCCCTAATATTCAATTAACATCCGAG 5279
DB 5236 GCTGCTGTAAAAAGCGGTTTCAGTTTAAAGTGGTCCCAAAATACCGGTATTACACAAAG 5295
QY 5280 GATGCTGAGAAGCACCGCTGGCAAGTGTTAGTGAATATCAAGGAGTGGCTATGGAAGA 5339
DB 5296 GATGAGAAAGAAACATATGGCAGCTCTTAGTAATATTAATAATGTAATAATGAGTGG 5355
QY 5340 GGATGCTGTCTTTATCTTTGGAGTTCGTTTCAATTTGTGTAGTACATAAAAAATAATGTA 5399
DB 5356 GGCTACTGCGCTTTGTCATTAAGAAATTTGTGTCTGTGTATGTTTATAAAAAATAATATA 5415
QY 5400 AGAAAAGGTTTGAAGGAAAGTATTTTGTAGTGTGACAGAGCGGCTCGCCAAATGTAATCTACT 5459

RESULT 5
AAZ20642
ID AAZ20642 standard; RNA; 6395 BP.
XX
AC AAZ20642;

DB 5416 AAAATGGGTTTGAGGAGAAAGTAACGAGTGTGAAACGATGGAGGACCCATGGAACCTTCG 5475
QY 5460 GAAAAGGTTGTTGAGGAGTTCGTGATCAAGTACCAATGGCTGTGAAACTCGAAAAGGTT 5519
DB 5476 GAAAGGTTGTTGATGAGTTCATGGAGATGTTCCAAATGTCGTTAGACTCGCAAGTTT 5535
QY 5520 CCGGAAAAACAAAAAGAAATGGTAGTTAATATGTTAAATAAT- 5561
DB 5536 CGAACCAATCTCAAAAAGAGGTCGGAATAATAATAATAATTTAGGTAAAGGGCGTTCA 5595
QY 5562 -----AAGAAAAATAAATAACAGTGGTAAGAGGTTTTTAAATTTGAGAAATGAGGAT 5615
DB 5596 GCGGAGAGCCCTTAAACCAAAAAGTTTGTATGAAGTTTAAAAAGAGCTTTAAATAATTTGAT 5655
QY 5616 AATGTAAGTGTGATGACGAGTCTATCGGTCATCGAGTACGTTTAAATCAATATGCTTATA 5675
DB 5656 AAGAAATTAAGCCGAGACGTCGGTCGCGAATCTCGATTCGATTA---AATATGCTTACT 5712
QY 5676 CAATCAACTCTCCGAGCCAAATTTGTTTACTTTATCTTCCGCTTACGAGATCCTGTGCAGC 5735
DB 5713 CAATCACTTCTCCATCGCAATTTGTTGTTTGTCACTCTGTATGGCTGACCTATAGAAT 5772
QY 5736 TGATCAATCTGTGTATCAAAATGCAATGGGTAAACCGATTTCAAACGCAACAAGCTAGACAA 5795
DB 5773 TGTAAACGTTTGTACAAATTCGTTAGGTAAACCGATTTCAAACACAGCAAGCAAGACTA 5832
QY 5796 CAGTCCAAACAGCAATTTGCGGATGCTCGAAACCTGTGCTAGTATGACAGTGAGATTTTC 5855
DB 5833 CTGTTCAAACAGAGTTTCAAGCGAGGTGTGGAACCTTTCCCTCAGAGCACCGTCAGATTTTC 5892
QY 5856 CTGCATCGGATTTCTATGTGTATAGATATAATTCGACGCTTGTATCGGTTGATCAGCGGT 5915
DB 5893 CTGGCGATGTTTATAAGGTGTACAGGTACAAATGCAATGCAATTTTATAGTCTCTAACTACTCGGT 5952
QY 5916 TATTAATAGCTTTGATGATGATAAGATAAGATAAGAGTTGATATAATCAACCCGACCGA 5975
DB 5953 TGCTGGGGGCTTTTGTATCTAGGAATAGAAATAATCGAAGTAGAATAACCCAGCAGAGTCCGA 6012
QY 5976 ATACTACTGAATCGTTTAAACGCTACAGAGGTAGACGATGCTACTGTAGCTATAAGGG 6035
DB 6013 CAACAGCTGAAAAGTTAGATGCTACCCGACGGGTAGACGACGCTACGGTTGCAATTCGGT 6072
QY 6036 CTTCAATCAATTAATTTGGCTAATGAACCTGTTCTGGAACTGGCATGTTCAATCAAGCAG 6095
DB 6073 CTGCTATAAAATAATTTAGTTAATGAACCTAGTAAGAGGTACTGGACTGTACAATCAGAATA 6132
QY 6096 GCTTTGAGACTGTAGTGGACTTGTCTGGACCAACCTCCGCGCTACTTTAGCTATTGTTGT 6155
DB 6133 CTTTGTGAAGTATGTCGTGGGTGGTCTGGACCTCTGCACCTGCACTCTTAAATGCAATGGT 6192
QY 6156 GAGATTTCTTAAATTAAGTCTGTAAGACTTAAATTTACGGGTGGCTGATACCAAAATC 6215
DB 6193 GCTGAATATAAA-----GTTTGTGTTTCTTAAACACACAGTGGTACGTAACG 6244
QY 6216 AGCAGTGGTGTGTCGCTCACTTAAATATAACGATTTGTTCATATCTGGAATCCCAAGTTAAA 6275
DB 6245 TACAGTGTGTTTCTCCCTCCACTTAAATCGAAGGTTAGTGTCTTGGAGCGCGGAGTAAAC 6304
QY 6276 CCATGTGATGTTGATGATGTTGTTGATGCGGTAAACATCGGAGAGGTTCGAATCTCTCCC 6335
DB 6305 ATATATGGTTCATATATGTCGCTAGGACGCTAATAAAAAAGCGAGGATTCGAATCTCCCG 6364
QY 6336 TAACCGCGGTAGCGGCCCA 6355
DB 6365 GAAACCCCGTTGGGGCCCA 6384

1338 GAAUCGUAUUCGAGGUAUAUUAACGUGUGACAGCGAGGUCGAAUGGAGUG 1397
1400 GATAAGCAATCTTCAACCTTGTCAATGACTTCTTCTGCGAGACTAAGCTGGCTCG 1459
1398 GACAAUUCUUUGUAUUAUUCUUGGAGGUAUUAUUAUUAUUAUUAUUAUUAUUA 1457
1460 CTTCAAGACGATATAGTAATGGAAGTTTCGGTCTTGGTAAGACCACTTCTGAAT 1519
1458 CUAAAGGAGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1517
1520 ATTTGGGATGAGGTGGGCAAAATTTTGGAAACGTTTTCCCACTATCAAGAGAGATTG 1579
1518 GUGUGGAGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1577
1580 GTGACGAGAAATCTGATGTAAGTGAATGCTCTGAAGATCAAGATCCAGATCTG 1639
1578 UUGAACAGGAACUUUAUCAGAGUGGACGCGGACGCAUUAUUAUUAUUAUUAUUA 1637
1640 TATGTCAATGGAAGACAGAGTTCTGTAGTGAATACCAAGTCTGAGAGTTACCGCAT 1699
1638 UAUGUACCUUCCACGACAGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1697
1700 CTAGATATCAAGAGGACTTAGAAGACTGAGCAAAATGTACGACGCTTATCAGAAAT 1759
1698 CUUGACAUUAGGAAGAGGAGAAACGGAAGUGAUGUAUUAUUAUUAUUAUUAUUA 1757
1760 TCTATCCTTAAGGTGCTGATTAATTTGATATCGGAAGTTCAAGACATGTGCAAGCT 1819
1758 UCGUGUUAAGGAGUCUGACAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1817
1820 TTAGATTTAGTCTGATGTGACGACGAGTAATCTGTGAGTGGCGGAGAGATAGAAGC 1879
1818 UUGGAAGUUGACCAUAGACGACGAGGAGUUAUUAUUAUUAUUAUUAUUAUUAUUA 1877
1880 GGTTAATCTTATCTTTGATTAAGCAACCGAGGAGAAATGTGGCTTAAGCTCT-----T 1933
1878 GGUUGACUCUCAUUUAAGACGACUACUGAGGGAUUGUGCGUAGCUUUAUUAUUAUUA 1937
1934 AAAAGCAGGCTCTGAGCGCTGTGTATCTTGAACGACATCCGAGAGGTGAACGTA 1993
1938 CAAGAGAGGCUUACAGAGGUGCAUUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1997
1994 AATAAATTTCTATTCTGAGAAAGGAGATTCCTGTGTGTCGAGAAAGTCAATGGTTTG 2053
1998 AUGAGGUGUUCGAGGCCAGAGAGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 2057
2054 ACGAATGTAACTTAGACACACAGAGTTGGAGTCCCTCAACGATTTCCATAGGCTTGC 2113
2058 GAGUGGUCUUAUUAUUAAGAACGAGGAGUAGAGUCUUAUUAUUAUUAUUAUUAUUA 2117
2114 GTGGATAGTGTATTAACAAGCAATGGCATCGTGTGTCTACACTGGCTCACTCAAGTT 2173
2118 GGAUUCGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2177
2174 CAACAAATGAGAACTATATGGAAGTTTGGAGCTTCGTGTGCGCCACTGTATCAAT 2233
2178 CAGCAAAUAGAAACUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2237
2234 CTAATGCAATCTAATAAGATGAGTGGGTATGATCTCTGATTCAGGGAGAAAGTTGTT 2293
2238 CUCGUCAAGAUCCUCAAAGUAUACGUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2297
2294 GTTTGGGATGCTACTTTGAAAGAGTGGCTCTCAACCTCGCGCAAGGCTATTCTATGG 2353
2298 GUCUUGGAGUUGCAUUAAGGAAGUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 2357
2354 GGAGTTGTCTGATTTAAGGGGAAATGTTTACTGCACTTCTATCTTATGAAGAGAT 2413
2358 GUGUUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2417
2414 AGAATGCTAGTGAAGCACTGGAGGAGGTGCTGTATCATCTCTATCAATGTTATAT 2473
2418 GGUGUGGAGACUUGCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2477

2474 TCTGATATTGCAAAAGCTCAAAATCTGAGGAAACAAATGAGAGACGGTGAACCCACGAA 2533
2478 UCCGACUUGGCAAAACUCAGAAACUCUGCGCAGACUGCUUCAAACGAGAACCCGACUC 2537
2534 CTTACTCAAAAGATGTGATCTTGTGATGGGTGCTGGTGTGAAAGTCAAAAGGAGAT 2593
2538 AGUACGCAAGGUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2597
2594 TTTGAAAGATTTGATCTTGTGAGAGATTTGATCTTGTGCTGGAACAAAGCTGCTGCT 2653
2598 CUUUCAGGGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2657
2654 ATGATCAGAAAGAGGCTTAATCTCTGGACTGATAAGAGCCACAATGGAACAATGTGAGA 2713
2658 AUGAUCAGAAAGCUGCGAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2717
2714 ACGGTAGATTCATCTTCTTAATGAT-----CCAAACCGCGATCACAAAGAGGCTT 2764
2718 ACCGUGAUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2777
2765 TTTATTGATGAGGTTGATGCTGCACACCGGTTGTGTTAACTTCTGCTGCTTATCTCT 2824
2778 UUCAUUAAGAGGUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2837
2825 GGTTCGACATCGCATACATTTACGAGATACACAGCAGATTTCTTTTCAATTAACAGATT 2884
2838 UUGUGGGAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2897
2885 CAGAAATTTCCGATTCCTCAAAATTTTGAAGCTGGAAGTGAAGTGAAGTGAAGTGA 2944
2898 UCAGAUUCCCGUACCCCGCCCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2957
2945 AGGACCACTGAGATGCCAGGTGATGATGATTTTCTTCAATCAAGTACGAAGA 3004
2958 AGAACUACUCCCGUUGUCCAGCCGAGUGUACACAUUAUUAUUAUUAUUAUUAUUA 3017
3005 GCGGTGACAAACCACTTCAACTGTACACGATCGGTCTCATCTGAGATGATAGCGGTAG 3064
3018 UUUUGAUGACACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3077
3065 GAGTACTAAACAGTGTTCCTCAACCACTAAAGGGAATTTGTAATCTTCACTCAGCT 3124
3078 GCGGUGAACAUCCGACUCAAACCCUUGCAUGCAAGACUCCUUAUUAUUAUUAUUA 3137
3125 GATAAATTTGATTTAGAGAGAGGCTATAAGAAATGTGAACACACCGTTCATCAGATCCA 3184
3138 GAUAAAGAGCUCUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3197
3185 GGAGAAACCTTTGAAGATGTGTGCTGATGATGAGGCAACTCCACTGATCTGTATT 3244
3198 GCGGAGACAUACUCUGAUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3257
3245 TCCAAGTCTTCCCGCATGTTCTAGTCTGCTCTGACTAGACACACAAAGAGCTTCAAAAT 3304
3258 GCAGGAGACAGCCCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3317
3305 TACACGCTAGTGTAGATCTTTAGTACAGATAATTTAGTGAATTTGTTCTTTTAAGCTCC 3364
3318 UACACUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3377
3365 TTTCTTTTGAAGATGATATGATGAGAGAGGATGATGATGATGATGATGATGATGAT 3424
3378 UACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3437
3425 GTGTTCAAGGTCATAATCTTTTGTGGCAACACCTTAATCAGGAGACTTCCAGATCTA 3484
3438 GUGUUAAGGUGUCCAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3497
3485 CAGTTCTTATTACGATGATGCTCTCTGTAATAGTACTTATCTTAAACAGTATGATGCT 3544
3498 CAGUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3557

Db 2838 UUGCGGAAAUUGCAUUAUUGCGGAGACACACAGCAGAUUCCAUCAUCAAUAUAGUU 2897
Qy 2885 CAGATTTCCTGATCCCAACATTTGAGAGCTGCAAGTGATGAAGTTGAGATGAGG 2944
Db 2898 UCAGAUUCCCGUACCCCGCCCAUUGGCAUUAUUGGAGAUUGACGAGUGGAGACACGC 2957
Qy 2945 AGGACACACAGTGGCGGAGTGGATGTGAATTTTCTCAATCGAAGTACGAAGGA 3004
Db 2958 AGAACUACUCCUGUUGCCAGCGCAUGUACACAUUAUCUGAAACAGGAGAUUAGGGGC 3017
Qy 3005 GCGGTGACACCACTTCACTTCAACGATCGATCGGTCTCATCTGAGATGATAGGCGGTAG 3064
Db 3018 UUGUCAUGAGACACUUCUGGUUAAAAGUGUGUUCGCGAGGAUGGUGCGCGGAGCC 3077
Qy 3065 GGAGTACTAAACAGTGTTCACAAACACCTATAAAGGGAATAATGTAACTTTCACTCAGGCT 3124
Db 3078 GCGGUAUACAUCGGAUCCUAAAACCCUUGCAUGGCAAGAUCCUGACUUUAUACCAUUG 3137
Qy 3125 GATAAATTGAGTTAGAGGAGAGGCTATAGAATGTGAACACCGTTTCATGAGATCCAA 3184
Db 3138 GAUAAAGAGCUCUGCUUUAAGAGGUAUUCAGAUUGUACACUGUGCAUGAAGUCNA 3197
Qy 3185 GGAGAAACCTTTGAAGATGTGCTGTGTCAGATTGACGGCAACCTCACTGACTCTGATT 3244
Db 3198 GCGGAGACAUACUCUGAUGUUUACUAGUUAAGGUUAACCCUACACAGUCUCCAUCAU 3257
Qy 3245 TCCAGTCTTCCCGCATGTTCTAGTCGCTCTGACTAGACACACAAAGAGCTTCAAATAT 3304
Db 3258 GAGAGACAGCCCAUUGUUUGUGCAUUGUACAGGCCACACUGUUGCUCUAGUAC 3317
Qy 3305 TACCCGTAGTTAGATCTTTTAGTACAGATAATAGTAGATTGTCTTTTAAAGTCC 3364
Db 3318 UACACUGUUAUGAUGUCCUUAUAGUUAUUAUAGAUUUAUAGAGAAACUUAAGCUCG 3377
Qy 3365 TTCTTTTGAATATGATGTAGTACAGAGTGTAGATGATGATGATGATGATGATGATG 3424
Db 3378 UACUUGUUAUGAUGUUAAGUUGCAGGAGGAAACAAUAGCAUUAUACAGAUUGACUG 3437
Qy 3425 GTGTTCAAAGTCTAATCTCTTTGTGGCAACACCTAATCAGGAGACTTTCCAGATCTA 3484
Db 3438 GUGUCAAAGGUUCCNAUUCUUUUGUGCAGCGCCCAAGACUGUGAUUAUUCUUAUG 3497
Qy 3485 CAGTTCTATTACGATGTATGCTCTCCCTGGTAAATAGTATATCTTAAACAGTATGATGCT 3544
Db 3498 CAGUUUAUUAUGAUAAGUGUCUCCAGGCAACAGCACCAUGAUGAAUUAUUAUGU 3557
Qy 3545 GTTACCATGAGTTACGTGATATAGTCTTAATGTGAAGGATGTGTTCTGATTTTCC 3604
Db 3558 GUUACCAUGAGGUAGACUGACAUUUAUUGAUGUCAAAGAUUGCAUUAUGGAUUGUCU 3617
Qy 3605 AAAAGTATTCGATGCCAAGGAGGTGAACCATGTCTAGAGCCAGTTTTCGCTACCGCG 3664
Db 3618 AAGUCUUGUGCGCCUAGGAUCAAUUAACAAACACAUUAUACCUAUGGUACGACGCG 3677
Qy 3665 GCGGAACCGCGAAGGCTCGAGACTCTCGAAATCTGTTGCAATGATTAAGAAAT 3724
Db 3678 GCAGAAUUGCCAGCGCAGACUGGACUUAUUGGAAAUUAUAGUGCGCAUGAUAUAAAGAA 3737
Qy 3725 TTCAACGCCACACCTGACGGGAGGATTTGATTTGAGAGACACCGCATCTGTTGTAGTA 3784
Db 3738 UUAACGCCACCGAGUUGUGCAUUAUGAUUAUUGAUAUUAUUGCAUUCUUGGUUGUA 3797
Qy 3785 GATAAGTTTTTATAGTATTTTATAAAGAAAAATATACAAAAAATATGCTGGA 3844
Db 3798 GAUAAUUUUUAUUAUUGUUAUUAAGAAAAAAGAAAAACCAAAUAAAAAUGUUUCU 3857
Qy 3845 GTGATGACGAAGATTCAATGATGATGATGTTGGAAAAACAGGAAGAACTACTATTGGAC 3904
Db 3858 UUGUACAGUAGAGUGUCUCUAUAGAUUGGUUUAAGAAACGAGGAACAGGTAAACAAUAGGC 3917
Qy 3905 GACTTGGTAACTTACAAATTTTACAGATCTCGCGGCGCATCGATCAGTACAGACACATGATC 3964
Db 3918 CAGCUCGCAUUAUUGAUUUUGGUAUUGGCAUUGCCAGCAGUUGAUCAGUACACAGACAUU 3977

Qy 3965 AAGGCTCAACCAAAACAGAAATTGGACCTTTCAATTAGAAATGAATACCCCTGCTCTGCAA 4024
Db 3978 AAAGCACAACCAAAACAGAAAGUUGGACACUUAUCCAAACGAGUAACCCGCGUUUGCAG 4037
Qy 4025 ACAATTGCTTACCACTTGAAGCAGATCAACGGTATTTTGGCGGTTTCTCA--GAGCTT 4081
Db 4038 ACGAUUGUUAUCCAUUCAAUAAAGAUCAUUAUUGCGCCCGUUGUUAUGAGUGCU 4097
Qy 4082 ACAAGGTTGCTGCTCGAGGCAATTTGATTTCTAAGAAATTTCTTTCTTTACTAGAAACT 4141
Db 4098 ACUAGCAUUAUUGGACAGUUGAUUUGGAGCAGAUUUUUUUUUAUUAUUAUUAUUAUUA 4157
Qy 4142 CCAGAACAGATTCAAGAAATTTTCTCGGATCTCGACTCGCAGCTTCTTATCGATGTGTTA 4201
Db 4158 CCAGCGCAGAUUGAGGAUUCUUGCGGAUUCUCGACAGUCUAGUGCGCGAUGAUGUCUUG 4217
Qy 4202 GAACTGATATTTCTAAGTATGATAAGTTCACAGAACGAGTTTCTTCTTCTGCTGTAGAGTAT 4261
Db 4218 GAGCUGGAUUAUCAAUUAACGAAAUUCAGAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4277
Qy 4262 GAAATATGAAAAAGATTGGGTCTCAATGAGTTTGTGGCCGAAGTGTGAAACAAAGGCAC 4321
Db 4278 GAGAUUGCGCAAGAUUGGUUUUAAGACUUCUUGGAGAGAUUUGGAAACAAAGGCAU 4337
Qy 4322 AGGAAAAACACTTTGAAGGATTACATTGCTGGAATCAAGACATGCTGTGTATCAAAAG 4381
Db 4338 AGAAAGACCAUCCUACAGGAUUAUACCGCAGUAUAAUAAACUUGCAUCUGGUAUCAAAGA 4397
Qy 4382 AAAAGCGGTGATGTGACTTACTTCTCGCAATCTGTTATAATAGCAGCTTGTCTTGGT 4441
Db 4398 AAGCUGGGAGCGUCACGACUUAUUGGAAACACUGUAUUAUUGCUGCAUUGUUGGCC 4457
Qy 4442 TCAATGTTTACCGATGAAAAAGGTCAATAAAGGTGCTTTTGTGGAGACGATTCGTTTTG 4501
Db 4458 UCGAUGCUCCGAGGAGAAAAUUAUAAUAAAGGAGCCUUUUGCGGAGCAUAGUCUGCUG 4517
Qy 4502 TATTTTCAAAGGTTTGGATTTCCCTGACATTCAGTCATGCTGCTAACTCATGTGGAT 4561
Db 4518 UACUUCCAAAGGUGUGAGUUUCGGAUGUGCAACUCGCGCAUUAUUAUUGGAAU 4577
Qy 4562 TTTGAGGCCAACTGTATAGAAAGGTACGTTACTTTTGTGTGTAGATACATCATACAC 4621
Db 4578 UUUUAGCAAAACUGUUUUAUAAACAGUUGAUAUUAUUGGAGAAUUAUUAUUAUUAUUA 4637
Qy 4622 CATGATAAGGAGCAATAGTGTATATGATCTTTTGAAGTGTATCTCAAACCTCGGGCA 4681
Db 4638 CACGACAGAGUUGCAUUGUUAUACGAUCCCUAAAGUUGAUCUCGAAACUUGUGUCU 4697
Qy 4682 AAACATATCAGGATTTATGATCATTAGAAAGTTAAGGTGCTTTTGTGCGATGTTGCT 4741
Db 4698 AAACAUAUAGGAUUGGAAACACUUGGAGGAGUUCAGAAUGUCUUUUGAUGUUGCU 4757
Qy 4742 TGTTCGCTCGGAAACTGTGTCTTAGCTTTCGCAAGTGAACGACGCTATCAAGGAGT 4801
Db 4758 GUUUCUU--GAACAUUGUGCGUAUUAACACAGAUUGGAGCAGCUGUAUGGAGGU 4814
Qy 4802 CATAAAACCGGATTTGATGTTGCTTTTAAATTTGTTTAAACAAATTTTTTGTGTAT 4861
Db 4815 CAUAAAGACCGCCUCCAGGUUGUUUUAUAAAGUUGGAGUUAUUAUUGUUGUUGCU 4874
Qy 4862 AAATTTTTTTTAAAGTCTTTTAAATGCTGTAGCTCAGAGATCTGTCAAAAT 4921
Db 4875 AAAGUUUUUUAAGAUUUUUUAUAGUGGCUUUAUAGUUUUUAAGGAAAGAAUUAUUAU 4934
Qy 4922 TAGCGAGTTCAATGATCTTTTCAAAACAGGATGAGATCTTCCGGCATTCATGACTAAGT 4981
Db 4935 CAUAGAUUUUAUCGACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4994
Qy 4982 CAAGAGTGTAGAAATATGATGCTGTGGACAAAGATTATGCTGTGTTAAGATGATGTTTC 5041
Db 4995 AAAGAGUGUUAUGUGUUCCAAAGUUAUAAUAAUUAUUAUUAUUAUUAUUAUUAUUAUUA 5054

QY 1934 AAAAGCAGCGGCTCTGAGCGCGTGTATGTCTTGAACCGACATCCGAAGAGGTGAACGTA 1993
Db 1938 CAAGAGAAGGCTTCAGAAAGGTGCTTTGGTAGTTACCTCAAGAGAAGTTGAAGAACCGTCC 1997
QY 1994 AATAAATTTTCTATTGCTGAGAAAGGAGATTGCCCTGTGTGCGAAGATCATGGTTTG 2053
Db 1998 ATGAAGGGTTCATGCGCCAGAGGAGATTCAATTAGCTGGTCTTCTGGAGATCATCCG 2057
QY 2054 ACGAATGCTAACTTAGAGCACAGGAGTTGGAGTCCCTCAACGATTTCCATAAGGCTTGC 2113
Db 2058 GAGTCGTCCTATTCTAAGAACGAGGAGATAGAGTCTTTAGAGCAGTTTCATATGSCAAG 2117
QY 2114 GTGGATGTGTATTACAAAGCAATGGCATCGGTGTCTACATCGGCTCACTCAAGTT 2173
Db 2118 GCGAATTCGTTAAATTCGTAAGCAGATGAGCTCGAATGTGTACACGGGTCGGAATTAAGTT 2177
QY 2174 CAACAAATGAAGAACTATGTGACAGTTTGGCAGCTTCTGTCGCGCACATGTATCAAAAT 2233
Db 2178 CAGCAATGAAGAACTTTATCGATAGCCTGTGTAGCATCACTATCTGTGCGGTGTGCAAT 2237
QY 2234 CTATGCAAGTCACTAAAGGATGAAGTCCGGTATGATTTCTGATTCAGGGGAGAAAGTTGGT 2293
Db 2238 CTCGTCAAGATCCTCAAGATACAGCTGCTATTGACCTTGAACCCGTCAAAAGTTTGA 2297
QY 2294 GTTTGGGATGTCACTTTGAAAGGTGGCTCCTCAAACTCGCGGCCAAAGGTCATTCATGG 2353
Db 2298 GTCTTGGATGTGTCACTAGGAAGTGGTTAATCAACCAACCGGCCAAGAGTCATGCAATGG 2357
QY 2354 GGAGTTGCTCGGATTACAAAGGGGAAATGTTTACTGCACTTCTATCTTATGAAGGAGAT 2413
Db 2358 GGTGTTGTTGAACCCACCGAGGAGGATATCATGTGGCGCTTTTGGAAATATGATGAGCAG 2417
QY 2414 AGAATGTGTACTGAGAGCGACTGAGGAGGGTGGCTGTATCATCTGATACAAATGGTATAT 2473
Db 2418 GGTGTGTGACATGCGATGATTGGAGAGAGTAGCTGTGAGCTCTGAGTCTGTTGTTTAT 2477
QY 2474 TCTGATATTGCAAGCTCAAAATCTGAGGAAACAAATGAGAGAGCGGTGAACCCCAAGAA 2533
Db 2478 TCCGACATGGCGAACTCAGAACTCTGCGCAGACTGCTTCGAAACCGAGAACCCGATGTC 2537
QY 2534 CCTACTGCAAGATGGTACTTGTGATGGGTGGCTGCTGTGTGGAACTACAAAGGAGAT 2593
Db 2538 AGTAGCGAAAGTTGTTCTTGTGACGAGTTCGCGGCTGTGGGAAACCAAGAAAT 2597
QY 2594 TTTGAAAGATTGATCTTTGATGAGATTGATCTTTGGTTCTCTGGAACAAAGCTGCTGCT 2653
Db 2598 CTTTCCAGGTTAAATTTTGTATGAAGATCTAAATTTTGTAGTACCTGGGAAGCAAGCCGGA 2657
QY 2654 ATGATCAGAAGAGGGCTAAATTCATCTGGACTGATGAAGAGCAATGGACAAATGTGAGA 2713
Db 2658 ATGATCAGAAGACGTGCGAAATTCCTCAGGGATTAATGTGGCCACAGAGGACAAACGTTAA 2717
QY 2714 ACGGTAGATTCACTTCTAATGCAT-----CCAAACCCGATCACAGAGGCTT 2764
Db 2718 ACCGTTGATCTTTTCATGATGAATTTTGGGAAACGACACGCTGTCAAGGAGTTA 2777
QY 2765 TTTATTGATGAAGGGTTGATGCTGCACACCGGTTGTGTAACTTCTCGTGTGTTATCTCT 2824
Db 2778 TTCATTGATGAAGGGTGTGATGTGCACTACTGTTGTGTAAATTTTCTTGTGGCGATGCA 2837
QY 2825 GGTTCGCAATCGCATACATTTACGAGATACACAGCAGATTCCTTTTCATTAACAGATT 2884
Db 2838 TTGTGCGAAATGCAATATGTTTACGAGACACACAGCAGATTCATATCAATAGAGTT 2897
QY 2885 CAGAATTTCCGATATCCCAACATTTTCAGAGCTGCAAGTGGATGAAGTGGATGAGG 2944
Db 2898 TCAGGATTTCCCGTATCCCGCCCATTTTTCGCAATTTGGAAGTTGAGAGGTGGAGACGC 2957
QY 2945 AGGACACACTGAGATGCCAGGTGATGTGAATTTTTTCTACAACTCGAAGTACCAAGGA 3004
Db 2958 AGAATCTCTCCGTTGTCCAGCCGATGTACACATATCTGAAACAGGAGATATGAGGGC 3017

QY 3005 CGGTGACCAACCACTTCACTGTACACGATCGGTCTCATCTGAGATGATAGGCGGTAAAG 3064
Db 3018 TTTGTCATGACACTTCTTCGTTTAAAGTCTGTTCGCGAGAGATGGTCGGCGGAGCC 3077
QY 3065 GGAGTACTAAACAGTGTTCACAAACCACTAAAGGGAATAATTGTAACTTTTCACTCAGGCT 3124
Db 3078 GCGGTGATCAATCGATCTCAAAACCTTGCATGGCAAGATCCTGACTTTTACCCAAATCG 3137
QY 3125 GATAAATTTGAGTTAGAGGAGAAAGGCTTATGAAGATGTGAACACCGTTTCATGAGATCCAA 3184
Db 3138 GATAAAGAGCTCTGCTTTCAGAGGGTATTCAGATGTTTCACTGTGTGATGAAGTGCAA 3197
QY 3185 GGAGAAACCTTTGAAGATGTGTCGTGATGAGCGCAATCCACTGATCTGATTT 3244
Db 3198 GCGGAGACATCTCTGATGTTTCACTAGTTAGTTTAAACCCCTACACCGAGTCTCCATCAT 3257
QY 3245 TCCAAGTCTTCCCGCATGTTCTAGTCCTCTGACTAGACACACAAAGAGCTTCAAAAT 3304
Db 3258 GCAGGAGACGCCACATGTTTGGTCGATTTGTCAGGCACACCTGTTGCTCAAGTAC 3317
QY 3305 TACACCGTGTGTTAGATTCCTTTTGTAGTACAGATAATTAGTGTGTTGCTTCTTTAAGTCC 3364
Db 3318 TACACTGTTGTTATGGATCCTTTAGTTAGTATCATTTAGAGATCTAGAGAAACTTAGCTCG 3377
QY 3365 TTCTTTTGTAGAAATGTATATGTTAGACGAGTAGTAGTACGATTTACAGATGGATGCA 3424
Db 3378 TACTGTTGTAGATATGTTAAGTGTGATGCGAACAATAATAGCAATTTACAGATTGACTCG 3437
QY 3425 GTGTTCAAAGTGCATAATCTCTTTGTGGCAACACCTTAATCAGGAGACTTTTCAGATCTA 3484
Db 3438 GTGTTCAAAGGTTCCATCTTTTGTGCGGCCAAAGACTGTTGTGATATTTCTGATATG 3497
QY 3485 CAGTTCTATTACGATGTATGCTCCTCGTGAATAGTACTATCTTAAACAGATGATGCT 3544
Db 3498 CAGTTTACTATGATAGTGTCTCCAGGCAACAGCACCATGATGAATAAATTTTGTATGCT 3557
QY 3545 GTTACCATGAGTTACGTGATAATAGTCTTAATGTGAGGATGTGTTCTTGTATTTTCC 3604
Db 3558 GTTACCATGAGTTGACTGACATTTCAATTGAATGTCAAGATTTGCATATTTGATATGCT 3617
QY 3605 AAAAGTATTCGATGCCAAGGAGGTGAACCATGTCTAGAGCCAGTTTGTGTAACCGG 3664
Db 3618 AAGTCTGTTGTGCGCTTAGGATCMAATCAACCACTAATACCTATGTTAGTACGACGCG 3677
QY 3665 GCGAAGCCGCAAGGGCTGACGAGTACTCGAAATCTGTTGCAATGATTAATAAGAAAT 3724
Db 3678 GCAGAAATGCCACGCGACACTGGACTATTGGAAATTTAGTGGCGATGATTAAGAAAGAAC 3737
QY 3725 TTCAACGCCACGACACTGACCGGGACGATTCACATTTGAGAGCACCGCATCTGTTGATGA 3784
Db 3738 TTTAACGCCACCGAGTTGTCTGGCATCATTTGATATTGAAATATCTGTCATCTTTAGTTGA 3797
QY 3785 GATAAGTTTTTTGATAGCTATTTTATAAAGAAAGAAATACAAAAAATATGCTGGA 3844
Db 3798 GATAAGTTTTTTGATAGTTATTTGCTTAAGAAAGAAAGAAACCAATATAAATGTTCT 3857
QY 3845 GTGATGACGAAGGATTCATCATGATGAGTGGTTGGAAAAAGAAAGAGTACTATTTGGAC 3904
Db 3858 TTGTTTCAGTAGAGTCTCTCAATAGATGGTTTAGAAAAAGCAGGAACAGTAACAATAGGC 3917
QY 3905 GACTTGGCTACTACAAATTTTACAGATCTGCGGCCATCGATCGATCAAGACACATGATC 3964
Db 3918 CAGCTCGCAGATTTTGTATTTTGTAGATTTGCGCAGATTTGATCAGTACAGACATGAT 3977
QY 3965 AAGGCTCAACCAAAACAGAAATTTGACCTTTCAATTCAAGATGAATACCTGCTCTGCA 4024
Db 3978 AAAGCAACCAACAGCAAAATTTGGACACTTCAATCCAAACGAGTACCCGGCTTTGCA 4037
QY 4025 ACAATTTGCTACCAATTCGAAGCAGATCAACCGTATTTTGGCCGGTTTCTCA---GAGCTT 4081
Db 4038 ACGATTGTGTACCAATTCAAAAAAGATCAATGCAATATTTGGCCCGTCTGTTAGTGAGCTT 4097
QY 4082 ACAAGTTGCTGCTCGAGGCATTTGATTTCTAAGAGATTTCTTTTCTTTTACTAGGAAACT 4141

Db 4098 ACTAGGCAATTAAGTGGAGTGTGATTCGAGCAGATTTTGTGTTTTCACAAAGAGACA 4157
Qy 4142 CCAGAAACAGATTCAGAAATTTTCTCGGATCTCGACTCGCAGTTCCTATGGATGTGTA 4201
Db 4158 CCAGCCAGATTTAGAGATTTCTTCGGAGATCTCGACAGTCATGTGCCAGTGTCTTG 4217
Qy 4202 GAACTGGATATTTCTAAGTATGATAAGTCACAGAACGAGTTTCAATGTGCTGTAGAGTAT 4261
Db 4218 GAGCTGGATATATCAAAATACAGCAAACTCAGATGAAATCCACTGTGCGAGTAGATAC 4277
Qy 4262 GAAATATGAAAAGATTTGGTCTCAATGAGTTTTTGGCCGAAGTGTGGAACAAGGGCAC 4321
Db 4278 GAGATCTGGCGAAGATTTGGTTTGAAGACTTCTTGGGAGAGTTTGGAAAACAAGGCAT 4337
Qy 4322 AGGAAAACAACTTTGAAGATTTACATCTCTGGAATCAAGACATGCTGTGGTATCAAAG 4381
Db 4338 AGAAGAGACACCTCAAGGATTTATACCGCAGGTATAAAACCTTGATCTGGTATCAAAGA 4397
Qy 4382 AAAAGCGGTGATGACTACTTTTCATCGCAATACCTGTTATATAGCAGCTTGTGGGT 4441
Db 4398 AAGAGCGGGACGTCACGACGTTTCATTTGGAACACCTGTGATCATTTGCTGCAATTTGGCC 4457
Qy 4442 TCAATGTTACCGATGGAAGGTCTAATAAGGTGCTTTTGTGAGAGAGATTCGCTTTTG 4501
Db 4458 TCGATGCTTCGATGGAGAAAATAATCAAAGGAGCCTTTTGGCGTCAAGATAGTCTGCTG 4517
Qy 4502 TATTTTCCAAAGGGTTTGGATTTCCCTGACATTCAGTCAATGCTTAATCTCATGTGGAT 4561
Db 4518 TACTTTCCAAAGGGTTGAGTTTCCGGATGTGCAACCTCCCGAATCTTAATGTGGAAT 4577
Qy 4562 TTTGAGGCCAAACTGTATAGAAAGGTACGGTTACTTTTGTGAGTATACATCATACAC 4621
Db 4578 TTTGAGCBAAACTGTTTAAAAAACAGTATGATACTTTTGGGAGATATATGATATACAT 4637
Qy 4622 CAGTATAGGGAGCAATAGTATATGATTCCTTTGAAAGTGTATCTCMAAATTCGGGCA 4681
Db 4638 CACGACAGAGATGCAATGTGATTAACGATCCCTAAAGTTGATCTCGAAACTTGGTGT 4697
Qy 4682 AACAATATCAAGGATTTATGATCACTTAGAAGAGTTAAGGGTCTCTTTGGCGATGTGCT 4741
Db 4698 AACAATATCAAGGATTTGGAACACTTGGAGAGTTTCAAGAGTCTCTTTGATGTGCT 4757
Qy 4742 TGTTCGCTCGAACTGTGTCTTAGGCTTTCCGAGCTGAAACGAGCTATCAAGAGTT 4801
Db 4758 GTTTCGTT--GAACAAATTGTGCGTATTACACACAGTTTGGACGCGCTGATGCGAGTT 4814
Qy 4802 CATAAAACCGGATTTGATGCTGTTGCTTTTAAATGTTTAAATGTTTAAACAAATTTTGTGAT 4861
Db 4815 CATAGACCGCCCTCCAGGTTTGTGTTTAAAGTCTGCTGAGAGTATTTGTCTGAT 4874
Qy 4862 AAATTTTATTTAGAACTTTGTTTAAATGCTGTAGTCTCAGAGATCTGTCAAAAT 4921
Db 4875 AAGTTCTTTTGAAGTTTGTATGATGCTCTAGTGTGTAAGGAAAGTGAATAT 4934
Qy 4922 TAGCGAGTTCAITGATCTTTTGAAGACAGGATGAGATCTTCGGGCAATCATCACTAAGT 4981
Db 4935 CAATGAGTTTATCGACTGACAAAATGAGAGAGATCTTACCGTCGATGTTTACCCCTGT 4994
Qy 4982 CAAGAGTTTAGAATATCACTGTGGACAAGATTTAGGCTGTGTAAGATGATGCTTTC 5041
Db 4995 AAGAGTGTATGTGTCTCAAAAGTTGATAAAAATAATGTTTCAATGAGATGATGCTATGTC 5054
Qy 5042 TGATGTAGATTTTACTTTAAAGGTGTTAAGTTAGTTAAGAAAGGTATGTGCTTAGCTGA 5101
Db 5055 AGAGGTGAACCTTCTTAAAGGATTTAAGCTTATGATAGTGAATGCTGTTTAGCCGG 5114
Qy 5102 TTTGGTGTGTCTGGGAGTGAATCTCCCGGATATCTCCCGTGGTGTGTCAGTGTGTTG 5161
Db 5115 TTTGGTGTGTCTGGGAGTGAATCTTGGCTGCAATTTGCAATTTGCAAGGAGGTGTGAGCGTGTG 5174
Qy 5162 TATTTAGATAGAGATGAAGAGGATGAAGNAGCAACGCTGGGTGCTATCACGCCCC 5221

Db 5175 TCTGTGGAACAAAGGATGGAAAGCGCAGAGGCCACTCTCGGATCTTTACTACACAGC 5234
Qy 5222 TGCTTGCAAAAAGAAATTTTCTTTTAAAGCTAATCCCTAAATTAATCAATAACATCCGAGGA 5281
Db 5235 AGCTGCAAGAAAGATTTCAAGTTCAAGGTGTTCCCAATTAATGCTATAACCAACCCAGGA 5294
Qy 5282 TGCTGAGAAAGCAACCGTGGCAAGTGTAGTGAATATCAAAGGAGTGGCTATGGAAGAGG 5341
Db 5295 CGCGATGAAAACGCTGGCAAGTTTGTAGTTAATTAAGAAATGTGAAGATGTCAAGCGGG 5354
Qy 5342 ATACTGTCTTATCTTTTGGAGTTTCTTCAATTTGTGTAGTACATAAAATTAATGTAAG 5401
Db 5355 TTTCTGTCGCTTCTCTGGAGTTTGTGTCGGTGTATTTGTATAGAAATAATATAA 5414
Qy 5402 AAAAGTTTGAAGGAAACGTAATTTTGAAGTGTGACAGACGCTCGCCAAATTTGAATCACTGA 5461
Db 5415 ATTAGTTTGAAGAGAGATTTACAAAGTGTGAGAGCGAGGCGCCATGGAATTTACAGA 5474
Qy 5462 AAAAGTTTGAAGAGTTCGTGGATGAAGTACCAATGGCTGTGAAACTCGGAAAGGTTTC- 5520
Db 5475 AGAAGTCTGTTGATGAGTTTCAAGAGATGTCCCTATGTGTCGATCAGGCTTGCAAGTTTCG 5534
Qy 5521 -----CGGAACAAACAAAGAAATGCTAGGTAATAATGTTAATAAT----- 5561
Db 5535 ATCTGAAACCCGAAACAAAGAGTGTCCGCAAGGGAATAATAGTAGTAATGATCGGTC 5594
Qy 5562 -----AAGAAAATAATAACAGTGTGTAAGAGGTTTAAAAATTGAGGAAATTTAGGA 5614
Db 5595 AGTCCGCAACAAAGATATAGAAATGTTAAGGATTTTGGAGGAAATGAGTTTAAAAAGAA 5654
Qy 5615 TAATGTAAGTGAATGACGAGT-----CTATCGCTCANTCGAGTACGTTTTTAATCAAT 5665
Db 5655 TAAATTAATCGATGATGATTCGGAGCTTCTGTCGCGAATCGGATTCGTTTTA---AAT 5711
Qy 5666 ATGCTTATACAACTCAACTCTCTCCGAGCCAAATTTGTTTACTTATCTTCGCTTTTACAGAT 5725
Db 5712 ATGCTTATACAGTATCACTACTACTCTCAGTTCTGTTCTTGTCTATCAGCGTGGGCCAC 5771
Qy 5726 CCGTGTGACGTGATCAATCTGTGTACAAATGCAATGGGTAAACCAAGTTTCAACGCAACAA 5785
Db 5772 CCAATAGAGTTAATTTAATTTATGATCTAATGCTTTAGGAAATCAGTTTCAACACAAACAA 5831
Qy 5786 GCTAGGACAAACAGTCCCAACAGCAATTTCCGGATGCTCGAAACCTGTGCTAGTATGACA 5845
Db 5832 GCTCGAACTGTCTGTTCAAGAACAAATTCAGTGAGGTGTGGAACCTTCCACCAAGTAAT 5891
Qy 5846 GTGAGATTTCTGTGATCGATTTCTATGTGTATAGATATAATTCGACGCTTGCATCCGTTG 5905
Db 5892 GTTAGTTTCCCTGACAGTGACTTTAAGGTGTACAGGTACATGCGTATTAGACCCGCTA 5951
Qy 5906 ATCAGCGGTTTATTAATAGCTTTGATATCTAGAAATAGAAATAATAGAGGTTGATATCAA 5965
Db 5952 GTCCACAGCACTGTTAGGTGCAATTCGACACTAGAAATAGAAATAATAGAAAGTTGAAAT 6011
Qy 5966 CCGGCAACGATATCTACTGAAATCGTTAAACCGGACTCAGAGGTGAGAGTCTACTGTA 6025
Db 6012 GCGAACCCCAACGACTGCGGCAACGTTAGATGCTCTCTAGAGTAGAGACCAACCGGTG 6071
Qy 6026 GCTATAAGGCTTCAATCAATAATTTGGCTAATGAATCTGTTTCTGTTGGAACCTGGCTGTT 6085
Db 6072 GCATAGAGGCGGATTAATAATTAATAGTAGAATTTGATCAGAGGAAACCGGATCTTAT 6131
Qy 6086 AATCAAGCAGGCTTTGAGACTGTAGTGAGACTTGTCTGGAACCAACTCCCGCTACTAG 6145
Db 6132 AATCGGAGCTCTTTCGAGAGCTCTTCTGGTTTGGTTTGGACCTCTGCTGCTGCTGCTG 6188
Qy 6146 CTATTTGTTGTGAGATTTCTTAAATAAAGTCTGCTGAAGACTTTAAATAATCAGGCTG 6205
Db 6189 -TGAGGTAGTCAAGATGATATAATAAACCGGATTTGTGTCGCTAATACCA-CGTGGTGG 6246
Qy 6206 TACCAAAATCAGCAGTGGTTGTTCTGCCATTTAAATAATAACGATTTGTTCATATCTGATCC 6265
Db 6247 TACGATACGCATAGTGTGTTTCCCTCCACTTAATTCGAAGGGTTGT-GTCTTGGATCCG 6305

1100 AAAACCTTGGCCATGTTCAAACCTGAAAGAGCAATCTTTAGAGACACGGCTTCGGTTAAC 1159
1098 AAGACTCTTGAATGCAACAGCAGAGAAATCCCTCTGAGGATTCATCATCAGTCAAT 1157
1160 TTTTGGTTCCTTAAGATGAAGACATGGTATAGTACCGCTGTTTGGAGGTTCTATTACC 1219
1158 TACTGGTTTCCAAAATGAGGATATGTCATCGTACCATTAATTCGACATTTCTTTGGAG 1217
1220 AGCAAAAAGATGCAAGAGTGAAGTCAATGTTAAATCGTGACTTCGGTTTACACAGTGTCT 1279
1218 ACTAGTAAGAGACGCGCAAGAAATCTTAGTGTCCAAGGATTTCTGTTTACAGTGTCT 1277
1280 AATCATATCAGAACATATCAAGCCAAAGGTTAACTTACCGAAAGTATTATCTTTCTG 1339
1278 AACCACATTCGAACATACACAGCGAAAGCTCTTACATACGCAAAATGTTTGTCTTTGTC 1337
1340 GAGTCTATAAGATCCCGGTGATTAATCAATGGTGTACTGTCTAGTCTGAATGGGATGTA 1399
1338 GAATCGATTCGATCGAGGTATCATTAACGCTGTACAGCGAGGTCCGAATGGGATGTG 1397
1400 GATAAGCAATTTCTCAACCTTGTCAATGACTTTCTTCTTCAGACTTAAGTGTGCTGG 1459
1398 GACAAATCTTGTCAATCTTGTCCATGACGTTTACCTGCATCTAAGCTTGCCTGTT 1457
1460 CTTCAAGACGATATAGTATGGAAGTTTCGGTCTTGATTAAGACCACTTCTCAACTT 1519
1458 CTAAGGATGACTTACTGATTAAGCAATTTAGTCTCGGTTCCGAACCGTGTGCCAGCAT 1517
1520 ATTTGGGATGAGTGGGCAAAATTTTGGAAAAGTTTCCCACTATCAAGAGAGATG 1579
1518 GTGTGGATGAGATTTCCGTGGGCTTTGGAAACGCAATTTCCCTCGTGAAGAGAGGCTC 1577
1580 GTGACGAGAAATTTCTGATGTAAGTGAATGCTCTCAAGATCAAGATCCAGATCTG 1639
1578 TTGAACAGGAAATTTATCAGAGTGGCAGCGCAATTAAGATCAGGTTGCTGATCTA 1637
1640 TATGTCATGGAAGACAGGTTTCGTAGCTGATACCAACCAAGTCTGAGGAGTTTACCGCAT 1699
1638 TATGTACCTTCCACACAGATTAAGTACTGATGACAGGCTCTGTGACATGCTCGG 1697
1700 CTAGATATCAAGAGGACTTAGAAGAGTGAAGCAATGTACGACCGGTTATCAGAAATTA 1759
1698 CTTGACATTAAGAAAGATGGAAGAAACGGAAGTATGATGATCAATGCACTTTTACAGAGTTA 1757
1760 TCTATCTTAAGGCTGCTGATATTTTCCGATATCGCAAGTTTCAAGACATGTGCAAGGCT 1819
1758 TCGGTGTTAAGGAGTCTGCAAAATTCGATGTTGATGTTTTCACGATGTGCCAATCT 1817
1820 TTAGATGTTAGTCTCTGATGTGGCAGCACGAGTAAATCGTTGCAAGTGGCCGAGAAATAGAAGC 1879
1818 TTGGAAGTTGACCAATGACGCGCAGCGAAGGTTTATAGTCGCGTCATGAGCAATGAGAGC 1877
1880 GGTTTAACTTTACTTTGATAAGCCAAACGAGAGAAATGTGGCTAAGGCTCT- - - - -T 1933
1878 GGTCTGACTCTCATAATTTGAACGACCTACTGAGGCGAATGTTGCGGTAGCTTTTACAGGAT 1937
1934 AAAAGCACGCGCTCTGAGGCGGTGTTATGTTCTTGAACCGACATCCGAAGAGGTGAACGTA 1993
1938 CAGAGAGGCTTCAGAGGTTCTTTGGTAGTTTACCTCAAGAGAGTTTGAAGAACGCTCC 1997
1994 AATAAATTTTCTATTGCTGAGAAGGAGATTCGCTGTGTGTGTCAGAAAGTCAATGGTTTG 2053
1998 ATGAAGGTTTCGATGGCCAGAGGAGATTAACAATTTAGTCTGTTGCTGGAGATCATCCG 2057
2054 ACNATGCTTAATTAGACACCGAGGTTTCGAGTCCCTCAACGATTTCCATTAAGGCTTGC 2113
2058 GAGTCGCTTATTTCAAGAACGAGGAGATAGAGTCTTTAGAGCAGTTTCAATATGGCAAG 2117
2114 GTGATAGTGTGATTAACAAGCAAAATGGCATCGGTTGTCTACACTGCTCACTCAAGTT 2173
2118 GCAGATTCGTTAATTCGTAAGCAGATGAGCTCGATTTGTGTACACCGGTCGGAATTAAGTT 2177

2174 CAACAAATGAAGAACTATGTGGACAGTTTGGCAGCTTCGTTGCCGCCACTGTATCAAAAT 2233
2178 CAGCAATGAAGAACTTTATCGATAGCTGTGTAGCATCACTATCTGTGCGGTGTCGAAT 2237
2234 CTATGCAAGTCACTAAAGGATGAAGTCGGGTATGATTTCTGATTTCCAGGAGAAAGTTGGT 2293
2238 CTCGTCAAGATCCCTCAAGATACAGCTGCTATTGACCTTGAAACCCCTCAAAAGTTTGGGA 2297
2294 GTTTGGGATGTCACTTTTGAAAAAGTGTCTCTCAAACTCGCGCCAAAGGTCATTCATG 2353
2298 GTCTTGGATGTCATCTAGGAAGTGTAAATCAAAACCAACGCGCAAGAGTCATGCAATG 2357
2354 GGAGTTGCTCGGATTAACAAGGGGAAATGTTTACTGCACCTTCTATCTTATGAAGAGGAT 2413
2358 GGTGTTGTTGAAACCCACGCGAGGAAGTATCATGTGGCGCTTTTGGAAATATGATGAGCAG 2417
2414 AGAATGCTGACTAGAGGACTGGAGGAGGTGCTGTATCATCTGATACATAATGTTAT 2473
2418 GGTGTGTGACATCGGATGATTTGAGAGAGTACTGTCTAGCTCTGAGTCTGTGTTGTTAT 2477
2474 TCTGATATTGCAAAAGCTCCAAAAATCTGAGGAAAAACAATGAGAGACGGTGAACCCCAAGAA 2533
2478 TCCGACATGCGCAAACTCAGAACTCTGCGCAGACTGCTTCGAAACGAGAAACCCGATGTC 2537
2534 CCTACTGCAAAAGATGTTACTTTGTGGATGGGTGCTGTTGTGGAAGTACAAAGAGAT 2593
2538 AGTAGCGCAAGGTTGTTCTTTGTGACGGAGTTCGGGCTGTGGGAAACCAAGAAAT 2597
2594 TTTGAAAGATTTGATCTTGATGAGGATTTGATCTTTGGTTCTCTGGAACAAGCTGTCTCT 2653
2598 CTTTCCAGGTTTAAATTTGATGAAGTCTAATTTTAGTACCTTGGAGCAAGCCGCGAA 2657
2654 ATGATCAGAAGAAGGCTTAATTCATCTGCACTGATTAAGAGCCACAATGGAACAATGTGAGA 2713
2658 ATGATCAGAAGACGTTGCGAAATTCCTCAGGGAATTTGTGGCCACGAGAGCAACGTTAAA 2717
2714 ACGGTAGATTCACITCTTAATGCTAT- - - - -CCAAAAACCGCATCACAAAGAGGCTT 2764
2718 ACCGTTGATCTTTCAATGATGAATTTTGGGAAAAAGCACACGCTGTCTGTTCAAGAGTTA 2777
2765 TTTATGATGAAGGTTGATGCTGCACACCGGTTGTGTTAACTTCTCTGGTCTTATCTCT 2824
2778 TTTATGATGAAGGTTGATGTTGCATACTGTTGTGTTAAATTTCTTGTGGCGATGCA 2837
2825 GGTTCGCACTCGCATACATTTACGGAGATACACGAGATTCCTTTTCAATTAACAGAGTT 2884
2838 TTGTCGGAATTCATATGTTTACGGAGACACACAGCAGATTTCCATACATCAATAGAGTT 2897
2885 CAGAAATTTCCCGTATCCCAACATTTTCGAAAGCTGCAAGTGGATGAAGTTGAGATGAGG 2944
2898 TCAGGATTTCCCGTATCCCGCCCATTTTGCAGAAATTTGGAAGTTGACGAGTTGGAGACACG 2957
2945 AGGACACACATGAGATGCCAGGTTGATGAATTTTTCCTACAATGGAAGTACGAAGGA 3004
2958 AGAATCTACTCTCGGTTCCAGCCGATGTACACATTAATCTGAAACAGGAGATATGAGGCG 3017
3005 GCGGTGACAAACCACTTTCAACTGTACACGATCGGTTCTCATCTGAGATGATAGCGGTAG 3064
3018 TTTGTGATGAGCACTTCTTCGGTTTAAAGTCTGTTTCGAGGAGATGTTGCGCGGAGCC 3077
3065 GGAGTACTTAAACAGTGTTCCTCAAAACCATTAAGGGAAGAAATTTGTAATTTTCACTCAGGCT 3124
3078 GCGGTGATCAATCCGATCTCAAAACCTTTGATGTCAGAGATCTCTGATTTTACCAATCG 3137
3125 GATAAATTTGAGTTAGAGAGAGGCTTAAAGATGTTGAACACCGTTTCAATGAGATCCAA 3184
3138 GATAAAGAGCTCTGCTTTCAAGAGGATTTTCAAGATGTTTCACTGTGATGATGAGTCA 3197
3185 GGAGAAACCTTTGAAGATGTTGCGTGTGTCAGATGAGCGCAATCTCACTGATCTGTGAT 3244
3198 GCGGAGACATCTCTGATGTTTCACTAGTTAGTTAAACCCCTACACCACTCTCCATCAT 3257
3245 TCCAAGTCTTCCCGCATGTTCTTAGTCTGCTGTGACTAGACACAAAGAGGCTTCAAAATAT 3304

Db 3258 GAGAGAGAGCCCATGTTTGGTCGCAATGTCAGGCAACCTGTTGCGTCAAGTAC 3317
Qy 3305 TACACCGTAGTGTAGATCCCTTTAGTACAGATAAATAGTGAATTTGCTCTTTTAAAGCTCC 3364
Db 3318 TACACTGTTGTTAGTGGATCCCTTTAGTGTAGTATCAATAGAGATCTAGAGAACTTAGCTCG 3377
Qy 3365 TTCCCTTTAGAAATGATATGTTAGTAGAGAGAGTAGATAGCAATTAACAGATGATGCA 3424
Db 3378 TACTTGTGTAGATATGATATAGGTCGATGAGGAAACAATAGCAATTAACAGATGACTCG 3437
Qy 3425 GTGTTCAAGAGTCAATAATCTCTTTGTCGCAACACCTAAATCAGGAGACTTTCCAGATCTA 3484
Db 3438 GTGTTCAAGAGTTCAAATCTTTTGTGCGCGCCAAAGACTGGTGATATTTCTGATAG 3497
Qy 3485 CAGTTCTATTAGATGATGCTCCCTGCTGATATGATCTATCTTTAAACAAGTATGATGCT 3544
Db 3498 CAGTTTCTACTATGATAAGTGTCTCCAGGCAACAGCACCATGATGAATAATTTTGATGCT 3557
Qy 3545 GTTACCATGAGGTTACGTGATATAGTCTTAATGTGAAGAGATTTGTTCTTGTATTTTCC 3604
Db 3558 GTTACCATGAGGTTGACTGACATTTCAATTGAATGTCAAAGATTGCATATTGGATATGCT 3617
Qy 3605 AAAAGTATTCGATGCCAAGAGGTGAACCATGCTAGAGCCAGTTTTCGTACCGG 3664
Db 3618 AAGTCTGTTGTCGCCCTAAGGATCAAAATCAACCACTAATACCTATGTTAGCAACGGCG 3677
Qy 3665 GCGGAACCGCAAGGGCTGCAGGACTACTCGAAAAATCTGGTTGCAATGANTAAAGAAAT 3724
Db 3678 GCGAAATGCCACGCCAGACTGGACTATGGBAAATTTAGTGGCGATGATTAAGGAAAC 3737
Qy 3725 TTCAACGCAACAGACTGACGCGGAGCATGTGACATGTAGAGACACCGCATCTGTTGATGA 3784
Db 3738 TTTAAGCGACCGGAGTTGCTGGCATCATGATATTGAATAATCTGTCATCTTTAGTTGTA 3797
Qy 3785 GATAAGTTTTTGTAGTCTATTTTAAAGAGAAATACAAAAAATATTCCTGGA 3844
Db 3798 GATAAGTTTTTGTAGTATTTGCTTTAAAGAAAAAAGAAACCAATTAATAAATGTTTCT 3857
Qy 3845 GTGATGACGAAGGATTCATGATGATGATGTTTGGAAAAACAGAAAAAGATGACTATTGGAC 3904
Db 3858 TTGTTCAGTAGAGTCTCTCAATAGATGTTAGAAAGCAGGACAGTAAACATAGGC 3917
Qy 3905 GACTTGGCTAACTACAAATTTTACAGATCTGCGGCGCATGATCAATCAAGACATGATC 3964
Db 3918 CAGCTCGCAGATTTGATTTTGTAGATTGCGCAGCATGATCACTACAGACACATGAT 3977
Qy 3965 AAGGCTCAACCAAAACAGAAATGACCTTCAATTCAGATGATACCTGCTCTGCAA 4024
Db 3978 AAAGCACAACCCAAAGCAAAATTTGACACTTCAATCCAAACGAGGATCCCGGCTTTGCAG 4037
Qy 4025 ACAATTTGCTACCATTCGAAGCAGATCAACGGTATTTTGGCGGTTTCTCA---GAGCTT 4081
Db 4038 ACGATTGTGTACCATTCAAAAGATCAATGCAATTTTGGCCCGTTGTTAGTGAGCTT 4097
Qy 4082 ACAAGTTGCTGCTCGAGCATTTGATCTAAGAAATTTCTTTTCTTACTAGGAAAACT 4141
Db 4098 ACTAGGCAATTAATCGACAGTGTGATTCGACAGATTTTGTGTTTTCACAAAGAAACA 4157
Qy 4142 CCAGAAACAGATTCAGAAATTTTCTCGGATCTCGACTCGCAGTTTCTTATGATGTGTTA 4201
Db 4158 CCAGCGCAGATTGAGGATTTCTTCGAGATCTCGACAGTCAATGTCGCGATGATGCTTTG 4217
Qy 4202 GAATCTGGATATTTCTAAGTATGATATGATCAAGACAGGTTTCAATTTGCTGTAGATAT 4261
Db 4218 GAGCTGGATATTAATAATACACAAATCTCAGATGAATTCACCTGTGAGTAGAATAC 4277
Qy 4262 GAAATATGAAAGATTTGGTCTCAATGAGTTTTTGGCGGAGTGTGGAACAAGGGCAC 4321
Db 4278 GAGATCTGCGAAGATTGGTTTTGAAGACTTCTTTGGGAGAGTTTGGAAACAAGGCAT 4337
Qy 4322 AGGAAAAACCTTTGAGGATTTACATTTGCTGGAATCAAGACATGCTGTGGTATCAAGG 4381

Db 4338 AGAAGACCCCTCAAGGATTATACCGCAGGTATAAAAAATTCGATCTGGTATCAAGA 4397
Qy 4382 AAAAGCGGTGATGACTACTTTTATCGGCAATACTGTTTATAATAGACAGCTTGTGGGT 4441
Db 4398 AAGAGCGGGACGTCAAGACGTTTCAATGGAAACACTGTGATCATTTGCTGCTATTTGGCC 4457
Qy 4442 TCAATGTTACCGATGGAAAGGTCATAAAAGTCTTTTCTGGAGAGCATTCCTGTTTG 4501
Db 4458 TCGATGCTTCCGATGGAGAAATTAATCAAGAGGCTTTTGGGTGAGCATGATCTGCTG 4517
Qy 4502 TATTTTCCAAAGGTTTGGATTTCCCTGACATTCAGTCAATGCTGTAATCTCATGTGGAA 4561
Db 4518 TACTTTCCAAAGGTTTGGAGTTTCCGATGTGCAACACTCCGCAATCTTATGTGGAA 4577
Qy 4562 TTTGAGGCCAACTGTATAGAAAGGTAACGGTTACTTTTGTGGTAGATACATATACAC 4621
Db 4578 TTTGAAGCAAACTGTTTAAAAAACAAGTATGATACCTTTTCCGGAAGATATGTAATACAT 4637
Qy 4622 CATGTAAGGAGCAATAGTGTATTTGATCTTTGAGTTGATCTCCAAACTTCGGGCA 4681
Db 4638 CACGACAGAGATGATGTTGTATTTACGATCCCTTAAAGTTGATCTCGAACTTGGTGT 4697
Qy 4682 AAACATATCAAGGATTTATGATCACTTAGAAGGTTTAAAGGTTCTTTTGTGCGATGTTGT 4741
Db 4698 AAACATATCAAGGATTTGGAACACTTTGAGGAGTTTCAAGGTTCTTTTGTGATGTTGT 4757
Qy 4742 TGTTCGCTCGAAACTGTTAGGCTTTCCGAGCTGAAACGACGATATCAAGAGGTT 4801
Db 4758 GTTTTCGTT---GAACAATTTGTCGTTATPACACAGATTTGGACGACCTGTATGGAGGTT 4814
Qy 4802 CATAAACCGGATGATGTTTGGTTTCTTTTAAATTTGTTTAAACAATTTTGTGTCAT 4861
Db 4815 CATAAAGCCGCCCTCCAGGTTTCTGTTTATAAAGTCTGGTGAAGTATTTGTGTCAT 4874
Qy 4862 AAATTTTATTTAGAACTTTGTTTAAATGGCTGTTAGTCTCAGAGATACCTGCAAAAT 4921
Db 4875 AAAGTTCTTTTGTAGAGTTTGTATAGATGGCTAGTTGTTAAAGGAAAGTGAATAT 4934
Qy 4922 TAGCGAGTTCAATGATCTTTCCAAACAGGATGAGATCTTCGGCATTCATGACTAAGGT 4981
Db 4935 CAATGAGTTTATCGACCTGCACAAAAATCGAGAGATCTTACCGTCGATGTTTACCCCTGT 4994
Qy 4982 CAAGAGTTTGAATATCGACTGTGGACAAAGTATGCTGTTAAGATGATGATGCTTTTC 5041
Db 4995 AAAGAGTGTATGTTTCCAAAGTTGATAAAATTAATGGTTTCAATGAGATGATGATGTC 5054
Qy 5042 TGATGTAGATTTACTTAAAGGTTTAAAGTTAGTTAAAGAAAGGATGTTGCTGCTAGCTGA 5101
Db 5055 AGAGTGAACCTTTTAAAGGATTTAAGCTTATTTGATAGTGAATAGTCTGTTTACCGG 5114
Qy 5102 TTTGATGTTCTGGGAGTGGAAATCTCCGGATAAATCTCCGCTGGTGGTGTGAGTGTG 5161
Db 5115 TTTGTTGCTACGCGGAGTGGAACTTCCCTGACAAATTCAGAGAGGTTGAGCGGTG 5174
Qy 5162 TATTTGATGAAGAGATGAAAGAGTGAAGAGCAACGCTGGGTGGTATACGCGCCC 5221
Db 5175 TCTGTGACAAAAAGGATGAAAGAGCGCAGCGCCACTCTCGGATCTTACTACACAGC 5234
Qy 5222 TGCTTGCAAAAGAAATTTTCTTTTAACTAATCCCTAATTTTCAATAAATCCGAGGA 5281
Db 5235 AGCTGCAAGAAAGAAATTTTCAAGTCTGTTCCCAATTTATGCTATAACCAACCCAGGA 5294
Qy 5282 TGCTGAGAGCAACCGCTGGCAAGTGTAGTGAATATCAAGAGGTTGGCTATGGAAGAAG 5341
Db 5295 CGCGATGAAACCGTCTGCGCAAGTTTGTAGTTAAATTAAGAAATGGAAGATGTCAGCGG 5354
Qy 5342 ATACTGCTCTTATCTTTTGGAGTCTGTTTCAATTTGTGTAGTACATAAAAAATATGTAG 5401
Db 5355 TTTCTGCTCGCTTTCTCTGGAGTTTGTCTCGGTGTATTTGTTTATAGAAAATTAATAA 5414
Qy 5402 AAAAGGTTTGAAGGAAACGATTTTGTAGTGTGACAGCGCTCGCCAAATTTGAATCACTGA 5461
Db 5415 ATTAGTTTGAAGAGAAAGATTTACAAACGTGAGAGCGGAGGCCCCCTTACAGA 5474

CC viruses, soluble proteins and peptides from plant sources. In order to
 CC isolate the bioactive species from the undesirable photosynthetic protein
 CC ribulose 1,5-diphosphate carboxylase (RuBisCo), the plant material is cut
 CC in a period of the light/dark cycle when the quantity of RuBisCo in the
 CC plant is at a minimum. The method is useful for obtaining a virus of
 CC interest. It is also useful for obtaining soluble recombinant or non-
 CC native proteins, such as active mammalian proteins, enzymes, vaccines,
 CC antibodies and peptides, from transgenic plants
 XX
 SQ Sequence 6439 BP; 1873 A; 1240 C; 1555 G; 0 T; 1771 U; 0 Other;

Query Match 38.6%; Score 2455.4; DB 4; Length 6439;
 Best Local Similarity 44.8%; Pred. No. 0;
 Matches 2772; Conservative 1147; Mismatches 2206; Indels 60; Gaps 8;

QY 20 ACAACAACAATTAACAAAAACAACATATTCAACAAACAACAACAACACATATGCGAC 79
 DB 18 ACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 77
 QY 80 ATACAATCTATTAATTAGCAACGCCCTTCTTGAAGCGTGGTGTAAACAACTCTCGTT 139
 DB 78 ACACAGACAGCAAC 137
 QY 140 AATGACCTTGCAAGAGGCGCATGTACGATACGCGCGTGGAGAGAAATTAAACGCCCGGAC 199
 DB 138 AAUGAUCUAGCAAGCGUGUCUUACGACACACAGCGGUGUAGAGUUAACGCGUGAC 197
 QY 200 CGTAGACCAAGTCACTTTTCCAAACTATTAGGAGAGCAACCGTCTTACTCTCC 259
 DB 198 CCGAGGCCACAGGUAACUUUUAAGUAAGCGAGGACGACGCUUAUUGCUACC 257
 QY 260 AACCGGTACCGGAGTTCCAGATTACTTTTATACTATCAAAATGCCGTACACAGTTTG 319
 DB 258 CGGCGUUAUCCAGAAUCCAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 317
 QY 320 GCTGAGGTTGAGAGCATTAAGATTGGAATCTGTACTGCTACAAGTTCCCTATGGATCG 379
 DB 318 GAGGUGGUAUGGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 377
 QY 380 CGCATATGATATAGTGGGAATCTTGCAGCACATTTTGCAGCAATTTTCAAGGCGAGGATACGTG 439
 DB 378 UUGCAUUAUGAUAUAGCGGGAUUAUUGCAUUGCAUUGCAUUGCAUUGCAUUGCAUUG 437
 QY 440 CATTGCTGATCCCAATCTGACATACGAGATATAATGAGCACGAGGACAAAGGAC 499
 DB 438 CACUGCUGAUGCCCAACUGGACGCUUCGAGACAUCAGCGGACGAGGCGCAGAAAGAC 497
 QY 500 TCAATTGAGATGATTTGTCAGATTGTCGTTCTTAACAGGTAATTCCTGAGTTTCAA 559
 DB 498 AGUAUUGAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 557
 QY 560 AGGAGGCTTTTAAACAGGTATGCAAGCTCCCAACGAGTCTGCTCTTAAATTTT 619
 DB 558 AAGGAGCAUUAUGACAGAUCCAGAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 617
 QY 620 CAGGATGTGCAATACATCCGCCAGAGATAGTGTGAGAGATACGCTGTTCTGTCAC 679
 DB 618 CAGACAAUGCGACAUCAGCGGAGCAGCAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 677
 QY 680 AGTTTGTATGATTTCTGTCATGATTTGGAGTCTGCTTAATATCTAAGNATATACAT 739
 DB 678 AGCAUUAUUGAUAUCCAGCGCAUUGUUGCGGCGGCAUUAUUAUUAUUAUUAUUAUUA 737
 QY 740 GTATGTTATGACGCTTCCATTTTGGCAGAAACATTAATTAAGACGAGCGAGGTTACG 799
 DB 738 AGGUGCAUUGCGCUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 797
 QY 800 CTTAATGAATAGGCGCACTTTTCAAAAGAGAGAGTGTATGTTCTTTTCTTCTGCT 859
 DB 798 UUGGACGAAAUCAAACGCGUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 857
 QY 860 GATGAAGTACTTTAAATATAGTCAATAAATAAATAATATCTTGCATTTAGTTAA 919

DB UCAGAGAGUACUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 917
 QY 920 TCTTACTTCTCTGCTTCTAGTAGATAGTATTACTTTTAGGAAATTTTACTCAGTAGGTT 979
 DB 918 ACUUAACUCCCGGCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 977
 QY 980 AATCTCTGTTTGTAAATTTTACCAAAATAGATACCTATATTCTCTACAAGAGTCTTAGA 1039
 DB 978 AAUACCUUGUUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1037
 QY 1040 CAAGTAGGTTGTATAGTATGATCAGTTCTATGAGGAGTGAAGAGCGCTTGTCTTACAG 1099
 DB 1038 CAUAAAAGUGUAGUAGUGAGCAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1097
 QY 1100 AAAACCTTGGCATTCTTCAACACTGAAGACCAATCTTTAGAGACACGCTTCGTTTAC 1159
 DB 1098 AAGACUUGCAUUGGCAACAGCGAGAGAUUCCUUAUUAUUAUUAUUAUUAUUAUUAUUA 1157
 QY 1160 TTTTGGTTCCCTAAAGATGAAGGACATGGTGTATGATAGTACCGCTGTTTGGAGGTTCTATACC 1219
 DB 1158 UACUGUUUCCAAAAUAGGAGUAUGGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1217
 QY 1220 AGCAAAAAGATGACAGGAGTGAAGTCAATGTTTAACTGTTGATCTGTTTACAGAGTGT 1279
 DB 1218 ACUAGUAAGAGGACGCGCAAGAAAGUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1277
 QY 1280 AATCATATCAGACATATCAAGCCAAAGCGTTAACTTACCAAGAGCTATTTCTTCGTG 1339
 DB 1278 AACCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1337
 QY 1340 GAGTCTTAAGATCCCGGTGATTAATCAATGGTGTACTGCTAGTGTCTGAATGGAGTGA 1399
 DB 1338 GAUCGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 1397
 QY 1400 GATTAAGCAATCTTCAACCTTGTCAATGATCTTCTTCTGAGATCAAGAGTGGCTGCG 1459
 DB 1398 GACAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1457
 QY 1460 CTTCAAGACGATATAGTAAATCGGAAAGTTTCGGTGTGATTAAGACACACTTCTGAACT 1519
 DB 1458 CUAAGGAGUACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1517
 QY 1520 ATTTGGGATGAGTGGGCAAAATTTTGGAAACGTTTTCCTTATCAAAAGAGAGATTG 1579
 DB 1518 GUGUGGAGUAGAUUUCGUGCGUUGGGAACCAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1577
 QY 1580 GTGAGCAGGAAATCTGGATGTAAGTGAGAACTCTCTGAGATCAAGATCCAGATCTG 1639
 DB 1578 UUGAACAGAAACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1637
 QY 1640 TATGTCATCGAAAGACAGGTTCTGATCTGAATACACCAAGTCTGAGGAGTTACCGCAT 1699
 DB 1638 UUGUGACCUUCCAGCAGACAGAUUAUGACUUAUUAUUAUUAUUAUUAUUAUUAUUAU 1697
 QY 1700 CTAGATATCAAGAGGACTTAGAGAAAGCTGAGCAAAATGTAAGCGCGGTTATCAAGATTA 1759
 DB 1698 CUUGACAUUAGAGAGAAUGAGAAACCGAAUGUUAUUAUUAUUAUUAUUAUUAUUAUUA 1757
 QY 1760 TCTATCTTAAGGCTGTGATAATTTTCGNATATCGGAATTCGAAAGATCTGCAAGCT 1819
 DB 1758 UCGUGUUAUAGGAGUUCUGAAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 1817
 QY 1820 TTAGATGTTAGTCTGATGTGGCAGCAGAGTAATTCGTTGAGTGGCGGCGAGAAATAGAGC 1879
 DB 1818 UUGGAUUGACCCAAUGACGCGCAGCGAAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1877
 QY 1880 GGTTTAACTCTTATCTTTTGTATAGCCAAACGAGAGAAATGTGGCTTAAGGCTCT-----T 1933
 DB 1878 GGUCUGACUCUCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1937
 QY 1934 AAAAGCAGCGCTCTGAGGCGGTGTTGTTCTTGAACCGACATCCGAGAGGTGAACGTA 1993
 DB 1938 CAAGAGAGGCUUCAGAGAGGUGCAUUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 1997

Db	3558	GUUACCAUGAGGUUGACUGACAUUUCAUGAUGUCAAAAGUUGCAUUAUUGGAUUGUCU	3617
Qy	3605	AAAAGTATTCCGATGCCAAAAGGAGTGAACCACTGTCTAGAGCCAGTTTTGGCTGACCGG	3664
Db	3618	AAGUCUUGUGCGGCCUUAAGGAUCAAACCAUUAUUAUUGGUACGAAAGCGCG	3677
Qy	3665	GCAGAACCGCAAGGGCTGCAGGACTACTCGAAAATCTGGTTCGAATGATTTAAAGAAAT	3724
Db	3678	GCAGAAUUGCCACGCCACAGACUGGACUAUUGGAAAAUUUAGUGGCGAUGAUUAAAAGCAAC	3737
Qy	3725	TTCAACGCCACGACCTGACGGGGACGATTTGACATTTGAGAGCACCGCATCTGTGTAGTA	3784
Db	3738	UUUAACGCACCCGAGUUGUCUGGCAUCAUUGAUUAUUGAAAAUAUCUGCAUCUUUAGUUGUA	3797
Qy	3785	GATAAGTTTTTTGATAGCTATTATTTATTAAGAAAAAGAAAAATACACAAAAATATTGCTGGA	3844
Db	3798	GAUUAAGUUUUUAGUAUGUUUUUUGCUUUAAGAAAAAAGAAAAACCAUAUAAAAUUGUUUCU	3857
Qy	3845	GTGATGCGAAGGATTCAATGATGAGATGTTGGAAAAACAGGAAGAAGTACTATTGGAC	3904
Db	3858	UUGUUCAGUAGAGUCUCUCAUAGAUUGUUAAGAAAAAGCAGGAACAGGUAAACAUAAGGC	3917
Qy	3905	GACTTGCTTACTACAATTTTACAGATCTGCCGCCATCGATCAGTACAGCACATGATC	3964
Db	3918	CAGCUCGCAUUAUUGAUUUUUGAUAUUUGCCAGCAGUUGAUCAGUACAGACAUGAUU	3977
Qy	3965	AAGSCTCAACCAAAAAAGAAATTTGACCTTTCAATTCCAGATGAATACCCCTGCTGTCAA	4024
Db	3978	AAAGCACACCCAGCAAAAAUUUGGACAUUUAUCCAAACGGAGUACCCGGCUUUUGCAG	4037
Qy	4025	ACAATTGCTACCAATTCGAAGCAGATCAACGGTATTTTGGCCGGTTTCT---CAGAGCTT	4081
Db	4038	ACGAUUGUGUACCAUUCAAAAAGAUCAUUGCAUAUUAUUGCCCGUUGUUGAGACUU	4097
Qy	4082	ACAAGTTGCTGTCGAGGCAATTTGATTCFAAGAGTTTCTTTTCTTTTACTAGAAAACT	4141
Db	4098	ACUAGGCAAUUACUGGACAGTUGUUAUUCGAGCAGAUUUUUGUUUUCACAAGAAAGACA	4157
Qy	4142	CCAGACAGATTCGAAGAAATTTTCTCCGATCTCGACTCGCAGCTTCCTATGGATGTGTTA	4201
Db	4158	CCAGCGCAGAUUGAGGAUUUUUCGAGAUUCUGCAGAUUGCGAUGGAUUGUCUUG	4217
Qy	4202	GAAC TGATATTTCTAAGTAGTAAGTACAGAACGAGTTTCATTGTGCTGTAGAGTAT	4261
Db	4218	GAGCUGGAUUAUCAAUAUACGACAAUUCAGAAUGAAUUCACUGCUGCAGUAGAUAC	4277
Qy	4262	GAATATGGAAGATTGGGTCTCAATGAGTTTTTTGGCCGAGTGTGTGAAAAACAGGGCAC	4321
Db	4278	GAGAUUCUGGCGAAGAUUUGGGUUUUGAAGACUUCUUGGGAGAAGUUGGAAAAACAGGGCAU	4337
Qy	4322	AGGAAAAACAATTGAAAGGATTACATTTCTGGATCAAGACATGCTGTGCTATCAAGG	4381
Db	4338	AGAAAGACCAUCCUACAGGAUUUAUCCGAGGUUAAAAACUGCAUCUGGUUAUCAAAGA	4397
Qy	4382	AAAAGCGGTGATGTGACTATTTTCATCGGCAATCTGTATTAATAGCAGCTTGTCTGGT	4441
Db	4398	AAGACGGGGACGUACAAGCUUAUUGAAAAACACUGUGAUCAUUGCUGCAUUGUUGGCC	4457
Qy	4442	TCAATGTTACCGATGGAAGAAAGGTCATAAAAAGGTGCTTTTGTGAGACGATTCGTTTTG	4501
Db	4458	UCGAUGCUCUCCGAUGGAGAAAAUAAUCAAAGGAGGCCUUUUGCGGUGACGAUAGUCUCUG	4517
Qy	4502	TATTTTCCAAAGGTTTGGATTTCCCTGCACATTCAGTCAATGCTATCTCATGTGGAAT	4561
Db	4518	UACUUUCCAAAGGGUUGUGAUUCCCGAUGUGCAACUCCCGCGAAUUCUUUUGUGAAU	4577
Qy	4562	TTTGAGGCCAACTGTATAGAAAGAGTACGGTTACTTTTGTGTAGATACATCATACAC	4621
Db	4578	UUUGAAGCAAAACUGUUUAAAAAACAUAUUGGAUUAUUUUGCGGAAGUAUUAUACAU	4637
Qy	4622	CATGATAAGGGAGCAATAGTGTATTATGATCTCTTTGAAGTTGATCTCCAACTTTGGGCA	4681

D	b	4638	CACGACAGAGGAUGCAUUGUUUAUACGAUCCCUAAAGUUGAUCUUGAAAUCUUGUGCU	4699
Q	y	4682	AAACATATCAAGGATTATGATCACTTAGAAGAGTTTAAGGGTGTCTTTGTGCAGATGTTGCT	4741
D	b	4698	AAACACAACGAUUGGGAACACUUGGAGGAGUTUCAGAAGGUCUCUUDUGAUGUUGUCU	4757
Q	y	4742	TGTTTCGCTCGGAAACTGTGGTGCTTTAGGCTTTCCGACGCTGNAACGACGCTATCAAGGAGGTT	4801
D	b	4758	GUTUCUGUAGAACAUUUGUGGUA---UUAACACACAGUUGGACGACGCGUAUUGGSGAGGU	4814
Q	y	4802	CATAAAACC CGCATGATGGTTCGTTTGCTTTTAAATGTTGTTAA CAAATTTTTTGTGTGAT	4861
D	b	4815	CAUAGACGCCCCUCCAGGUCUUGUUUAUAAAGCUCUGGUGAAGUAUUAUUGUCUAU	4874
Q	y	4862	AAATTTTATTTAGAACTTTGTTTTTAAATGGCTGTTAGTCTCAGAGATACTGTC AAAAT	4921
D	b	4875	AAAGUUCUUUUAAGAAGUUUUAUAGAGUGGUCUAGUUGUUUAAAGGAAAAAGUAAUAU	4934
Q	y	4922	TAGGAGTTCATTTGATCTTTCGAAAACAGNATGAGATATCTCCGGCAATTCATGATCAAGGT	4981
D	b	4935	CAAUGAGUUUAUCCGACCGUACAAAAAUGGAGAGAUAUCUUAACCGUCAUGUUUACCCCGU	4994
Q	y	4982	CAAGAGTGTTAGAATATCGACTCTGGACAAGATTAATGGCTGTTTAA GAATGATAGTCTTTC	5041
D	b	4995	AAAGAGUUUAUGUUGUCCAAAGUGAUA AAUAUGUUUAUGAGAAUGAGUAUGUAUUGUC	5054
Q	y	5042	TGATGTAGATTTACTTTAAAGTGTTTAAGTTAGTTTAAAGAAGGGTATGTGTCTTAGCTGA	5101
D	b	5055	AGAGGUAACCUUUAUAGGAGUAUAAGCUUAUUGAUGAUGGAUACGUCUGUUAAGCCGG	5114
Q	y	5102	TTTGGTAGTCTCTGGGAGTGGAAATCTCCCGGATAACTCCCGTGTTGGTGTCTAGTGTTTG	5161
D	b	5115	UUUGUGUCUACCGCGCAGUGGAACUUGCUCGACAAUUGCAGAGGAGGUGUGAGCGUGUG	5174
Q	y	5162	TATTTGTAGATAAGAGATGAAGAGGAGTAGGAAGCAACGCTCGGTGCGGTATCACGCCCC	5221
D	b	5175	UCUGUGGUA CA AAGGAUGGAAGAGCCGACGAGGCCACUCUCGGAU CUUAUACACAGC	5234
Q	y	5222	TGCTTGCAAAAAGAAATTTTCTTTTAAAGCTTAATCTCCCTAATTTTCAATAACATCCGAGGA	5281
D	b	5235	AGCUGCAAGAAGAAAUUUAUAGUUAAGGUCUUGUCCCAUUAUUGCUAUAACACCCAGGA	5294
Q	y	5282	TGCTGAGAAGCACCGGTGGCAAGTGTGTGTAATATCAAGGAGTGCGCTATGGAAGAGG	5341
D	b	5295	CGCGAUGAAAAACGUCUGGCAAGUUUUAUUAUUAUUAAGAAUUGAAGAUUGUACGCGGG	5354
Q	y	5342	ATACTGCTTTTATCTTTGAGTTCGTTTCCAATTTGTGTAGTACATAAAAATAATGTAAG	5401
D	b	5355	UUUCUGUCCGCUUUCUGGAGUUUGUGUCGGUGUGUAUUGUUUAUAGAAUAUAUAUA	5414
Q	y	5402	AAAAGTTTTCAGGGAACTATTTTGATGTGACAGACGGCTCCGCCAATTTGAATCTCACTGA	5461
D	b	5415	AUUAGUUUUGAGAGAGAAGAUUACAACGUGAGAGACGGAGGCCCAUGGAA CUUACAGA	5474
Q	y	5462	AAAGTTTGTGAGAGTTTCGTGGATGAAGTAPCAATGGCTGTGTAACCTCGAAAGGTTCC	5521
D	b	5475	AGAAUGCUUGAUGAGTUACUGGAAGAUGUCCCUAUGUCGAUCAGGCUUCGAAAGUUUCG	5534
Q	y	5522	GG-----AAACAAAAAGAAATGTTAGTGTATAATGTTAATAAT-----	5561
D	b	5535	AUCUCGAACCGGAAA AAGAGUGAUGUCCGCAAGGGGAAAUAUGUAUAUUAUGAUCGUC	5594
Q	y	5562	-----AAGAAAATAATAACAGTGTGTAAGAAGGGTTTTTAAAATTCAGGAAATTCAGGA	5614
D	b	5595	AGUCCCGAAACAGAA CUUAUAGAAUUAUAGGAUUUUGGAGGAUUGAUAUUAUAAAGAA	5654
Q	y	5615	TAATGTGAAGTGAACGAGT-----CTATCGGCTCATCGAGTACGTTTTTAATCAAT	5665
D	b	5655	UAAUUUAUUGAUGAUGAUUUCGAGGCUACUCUGUCGCGCAAUUCGGAUUCGUUUUA--AAU	5711
Q	y	5666	ATGCTTTATACATCAACTCTCCGAGCCAAATTTGTTTACTTATCTTCGCTTACGCAGAT	5725
D	b	5712	AUGUCUACAGUACUACUACUACUACUGUCUGUGUUUUGUUAUACAGUGGCGGCAC	5771


```
Db 5235 AGCUGCAAGAAAGAAUUCAGUUCAGGUCGUCCCAUUAUUGCUAAACACCCAGGA 5294
Qy 5282 TGCTGAGAGCACCCGTCGAGTGTAGTGAATATCAAGAGAGTGGCTATGGAAGAGG 5341
Db 5295 CGCGAUGAAACAGCUGCGCAAGUUAUUAUAAUAAUAGUAGUAGUAGUAGUAGUAG 5354
Qy 5342 ATACTGCTCTTATCTTTGGAGTTCGTTTCAATTTGTGTAGTACATATAAATAATGTAA 5401
Db 5355 UUCUGUCGCUUCUCUGCGAGUUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUGUG 5414
Qy 5402 AAAAGTTTGGAGGAAACGATTTTTCAGTGTGACAGACGCTCGCAATTTGAATCACTGA 5461
Db 5415 AUUAGUUUUGAGAGAAAGAUUACAAACUGAGAGACGGAGCGCCUUGGACUUAACAGA 5474
Qy 5462 AAAGTTTGTGAGGAGTTCGTCGATGAAGTATCAATGCTGTGTAACTCGAAAAGTTTCC 5521
Db 5475 AGAAGCUGUAGUAGUUAUGGAAGUUGCCUUAUGUGCAUGCAUGCGCUUCAAAGUUG 5534
Qy 5522 GG-----AAAAAAGAAAGTGTAGTGAATTAATTAATTAATTAATTAATTAAT 5561
Db 5535 AUCUGAACCGGAAAGAGUGUGUGCGCAAGAGGAAAGAAUAGUAGUAGUAGUAGUAG 5594
Qy 5562 -----AAGAAATATACAGTGTGAAGAGGTTTAAATTTGAGGAAATTTGAGA 5614
Db 5595 AGUGCCGAACAGAAACUAGAAUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 5654
Qy 5615 TAATTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 5665
Db 5655 UAAUUAUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAGUAG 5711
Qy 5666 ATGCTTATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCAATCA 5725
Db 5712 AUGCUUACAGUACUACUACUACUACUACUACUACUACUACUACUACUACUACUACU 5771
Qy 5726 CCGTGCAGCTGATCAATCTGTGTACAAATGATGCTGTGTGTGTGTGTGTGTGTGTGT 5785
Db 5772 CCAUAGAGUAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 5831
Qy 5786 GCTAGGCAACAGTCCAAACAGCAATTTGCGATGCTGTGTGTGTGTGTGTGTGTGTGT 5845
Db 5832 GCUCGAAUCUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 5891
Qy 5846 GTGAGATTCCTGATCGGATTTCTATGTGTATAGATATATATGATGATGATGATGATG 5905
Db 5892 GUUAGGUUCCUGACAGUGUACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 5951
Qy 5906 ATCAGCGGTTATTAATAGCTTTGATGATGATGATGATGATGATGATGATGATGAT 5965
Db 5952 GUCACAGCAGUUGUAGGUGCAUUCGACACUAGAAUUAUUAUUAUUAUUAUUAUUA 6011
Qy 5966 CCCGACCGAATATCTACTGAAATCGTTAAACGCACTCAGAGGTTAGACGATGCTACT 6025
Db 6012 GCGAACCCACGACUGCCGGAACAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 6071
Qy 6026 GCTATAAGGGTTCATCAATTAATTTGGCTAATGAACTGGTTCTGCGAATGCGATGTC 6085
Db 6072 GCGAUAAGGCGCGGAUAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 6131
Qy 6086 AATCAAGCAGGGTTGAGCTGCTAGTGTGCTGTGACCACTTCTGCGACCAATCTCG 6145
Db 6132 AAUCGAGCUCUUCUGAGAGCUCUUCUGGUGUUGUUGUUGUUGUUGUUGUUGUUG 6191
Qy 6146 CTATT 6150
Db 6192 CAAUU 6196
```

RESULT 13
AAZ20646
ID AAZ20646 standard; RNA; 6446 BP.
XX
AC AAZ20646;

```
XX 26-NOV-1999 (first entry)
DT
XX
DE TMV-based virus TMV861 coat protein read-through RNA sequence.
XX
XX TMV-based virus; tobacco mosaic virus; protein isolation; green juice;
KW virus isolation; fusion protein identification; ss.
XX
OS Tobacco mosaic virus.
XX
PN WO9946288-A2.
XX
PD 16-SEP-1999.
XX
PF 09-MAR-1999; 99WO-US005056.
XX
PR 10-MAR-1998; 98US-00037751.
XX
PA (BIOS-) BIOSOURCE TECHNOLOGIES INC.
XX
XX Garger SJ, Holtz RB, McCulloch MJ, Turpen TH;
XX WPI; 1999-561660/47.
XX
XX Obtaining protein, viruses and fusion proteins from plants, using non-
PT denaturing conditions.
XX
XX Disclosure; Page 55-58; 58pp; English.
XX
XX This sequence represents a tobacco mosaic virus (TMV) based virus
CC sequence identified using the method of the invention. The method is for
CC obtaining a soluble protein or peptide of interest from a plant,
CC comprises homogenising the plant to produce green juice, adjusting the pH
CC to less than or equal to 5.2, and heating the juice to a minimum of 45
CC degrees C. The juice is then centrifuged to produce a supernatant, and
CC the protein or peptide is purified from the supernatant. The method can
CC also be used for obtaining viruses and fusion proteins. The method is
CC especially useful for obtaining IL-1 to IL-10, EPO, G-CSF, GM-CSF, HP-
CC CSF, M-CSF, Factor VIII, Factor IX, tPA, receptors, neuropeptides, insulin,
CC antibodies, single-chain antibodies, enzymes, neurotrophic factors,
CC antigens, vaccines, peptide hormones, calcitonin, and human growth
CC hormone, or an antimicrobial peptide or protein from proteoglycans,
CC magainins, cecropins, melittins, indolicidins, defensins, beta-defensins,
CC cryptidins, clavalins, plant defensins, nisin and bacteriocins, all
CC produced by recombinant means. The new method is more efficient than the
CC prior art for isolating viruses, protein, and peptides. The method is
CC large-scale, and non-denaturing and solvent-limited. Prior art methods do
CC not isolate recombinant proteins, and do not allow fraction 2 proteins to
CC be ultrafiltrated
XX
XX Sequence 6446 BP; 1873 A; 1234 C; 1563 G; 0 T; 1776 U; 0 Other;
```

```
Query Match 38.5%; Score 2449.8; DB 2; Length 6446;
Best Local Similarity 44.8%; Pred. No. 0;
Matches 2771; Conservative 1146; Mismatches 2212; Indels 60; Gaps 8;

Qy 20 ACAACAAATTAATAAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 79
Db 18 ACCAACAAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACAACA 77
Qy 80 ATACATCTATATTAATAGCAACGCTTCTTGAAGCGTGTGTTAAACAACACTTCGTT 139
Db 78 ACACAGACAGCAUACCAUACGUCGUGGACGACGUCGAGGAAACAACUCCUUGGUC 137
Qy 140 AATGACCTTCGAGAGGCGCATGTACGATACGCGCGTGGAGAGATTTAAACGCGCGAC 199
Db 138 AUAUACUAGCAAGGCGGUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 197
Qy 200 CGTAGACCAAAAGGTCAACTTTTCCAAAACCTATTAGCGAAGAGCAAAACGCTTCTAGTCTCC 259
Db 198 CGCAGGCCCAAGGUGAACUUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 257
Qy 260 AACGGTACCCGGAGTTCAGATTACCTTTTATTAATCTCAAAATGCCGTACACAGTTTG 319
```


Qy	2474	TCTGATATTGCAAGAGCTTCCAAAATCTGAGGAAAACAATATGAGAGACGGTGTGAACCCCAAGAA	2533
Db	2478	UCCGACAUGCGCGAAACUCAGAAACUCUGCGCAGACUGUCUGGAAACGGAGAACCGCAUGUC	2537
Qy	2534	CCTACTCCAAGATGGTACTTGTGTGGATGGGGTGCTGTTGTGGAAAGTACAAAGGAGAT	2593
Db	2538	AGUAGCCGAAGGGUUGUUGUGUGGACGGAGUUCCGGGCUGUGGGAAAAACAAAGAAAUU	2597
Qy	2594	TTTTGAAAGATTTGATCTTTGATGAGGATTTGATCTTTGTTCTCTGAGAAAACAAGCTGTGCT	2653
Db	2598	CUUCCAGGGUUAUUUGAUGAAGAUUUAUUUAGUACUGCGGAAGCAAGCGCGAA	2657
Qy	2654	ATGATCAGAGAAGGGCTTAATTCATCTGGAATGATTAAGAGCACAATATGGACAATGTGAGA	2713
Db	2658	AUGAUCAGNAGACGUGCGNAUUCUCAGGGAUUAUUGUGGCCAGGAAGGACACGUAUAA	2717
Qy	2714	ACGGTAGATTCACTTCTTAATGTCATCCAAAACCGCGATCAC-----ACAAGAGGCTT	2764
Db	2718	ACCGUAGAUUUUUAUGAUGAUUUUGGAAAGACACACGUGUCAGUUCUACAGAGGUUA	2777
Qy	2765	TTTATTGATGAAGGGTTGATGCTGTCACACACGGTTGTGTTAACTTCCTGGTGCTTATCTCT	2824
Db	2778	UTUCAUUGAAGAGGGUUGAUGUGCAUACUGUUGUUAUUUUCUGUGGCGAUGUCA	2837
Qy	2825	GTTTGCACATCGCATACATTTACGGAGATACACAGCAGATTCCTTTCAATTAACAGATT	2884
Db	2838	UUGUGGAAAUUGCAUAUUGUUAACGGAGACACACAGCAGAUUCCAUACAUCAAUAGAGUU	2897
Qy	2885	CAGAATTTCCGTTATCCAAACATCTTTGAGAAGCTGCAAGTGGATGAAGTTTGAGATGAGG	2944
Db	2898	UCAGGAUUCGGUACCCCGCCAUUUUGCCAAUUGGAAAGACACACGUGUCAGUUCACAGAGUUUA	2957
Qy	2945	AGGACCACTGAGATGCCAGGTGATGTGAATTTTTTCTTACAATCGAAGTACGAAGGA	3004
Db	2958	AGAAUCUACUCUCGUGUCCAGCGAUGUACAACAUAUUGAACAAGAGAUUAGAGGC	3017
Qy	3005	CGGTGACACCACTTCAACTGTATACAGATCGGTCTCATCTGAGATGATAGCGGTAAG	3064
Db	3018	UUUGUCAUGAGCACUUCUCUGUUAAAAGUCUGUUGCGCAGGAGAUUGGCGCGAGCC	3077
Qy	3065	GGAGTACTAAACAGTGTTCCAAACCACTAAAAGGGAAAAATTGTAACTTTCACTCAGGCT	3124
Db	3078	GCCGUGAUAUCCGAUCUCAAACCCUUGCAUGGCAAGAUCCUGACUUAUCCCAUUCG	3137
Qy	3125	GATAAATTTGATGTAGAGAGAGAGGGCTATAAGATGTGAACACCGTTCTATGAGATCCAA	3184
Db	3138	GAUAAAAGACUCUCUGUUAAGAAGGGUUAUUCAGAUUACAUGUCUGAUGAUGAAGUGCAA	3197
Qy	3185	GGAGAACCTTTGAAGATGTGCTCGTTCAGATTGACGGCACTCCACATGACTCTGATT	3244
Db	3198	GCGGAGACAUAUCUCUGAUGUUUACUAGUUGUUUAAACCCUACACAGUCUCCAUCAU	3257
Qy	3245	TCCAAGTCTCCCGCATGTTCTAGTCGCTCTGACTAGACACACAAAGAGCTTCAAATAT	3304
Db	3258	GCAGGACAGCCCCACAUGUUUGUCGAUUGUCAAGGCACACCUUUCGCUCAAGUAC	3317
Qy	3305	TACACCGTAGTTAGATCTCTTTTAGTACAGATAATTAGTGATTGTCTTCTTTAAGTCC	3364
Db	3318	UACACUGUGUAUGGAUCCUUUAGUUAUUAUAGAUCAUUAAGAUUACUAGAGAAACUUAUGCUG	3377
Qy	3365	TTCCTTTTAGAAATGTATGTGTAGAGACAGGTAGTAGTAGCAATTTACAGATGGATGCA	3424
Db	3378	UACUUGUUAGUAUAGUUAAGUUCUAGCAGGAAACACAUAUAGCNAUUAACAUUGACUUG	3437
Qy	3425	GTGTTCAAAGGTCAATACTCTCTTTGTGGCAACACCTAAATCAGGAGACTTTTCCAGATCTA	3484
Db	3438	GUGUUCAAAGGUUCCAAUCUUUUUGUGCAGCGCCCAAGACUGGUGAUUUUUCUAUUG	3497
Qy	3485	CAGTTCATTATAGATGATGCTCCCTCGGTGAATAGTACTATACATTAACAAGATATGCT	3544
Db	3498	CAGUUUAUCUAGAUAAGAGUCUCCCGAGCAACAGCCCAUGAUGAUAUUAUUUGUUGCU	3557

was used to infect field-grown tobacco. The virus was then isolated from the tobacco plants by a novel process for isolating and purifying viruses, soluble proteins and peptides from plant sources. In order to isolate the bioactive species from the undesirable photosynthetic protein ribulose 1,5-diphosphate carboxylase (RuBisCo), the plant material is cut in a period of the light/dark cycle when the quantity of RuBisCo in the plant is at a minimum. The method is useful for obtaining a virus of interest. It is also useful for obtaining soluble recombinant or non-native proteins, such as active mammalian proteins, enzymes, vaccines, antibodies and peptides, from transgenic plants.

Sequence 6446 BP; 1873 A; 1234 C; 1563 G; 0 T; 1776 U; 0 Other;

. Query Match 38.5%; Score 2449.8; DB 4; Length 6446;

Best Local Similarity 44.8%; Pred. No. 0;

Matches 2771; Conservative 1146; Mismatches 2212; Indels 60; Gaps 8;

[illegible]

Db 1938 CAAGAGAGGCUUAGAGGUCUUGUGUUAUCCUACAGAGAGUUGAAGAACCGUCC 1997
Qy 1994 AATAAATTTCTATTGCTGAGAAAGGAGATTGCTGTGTGTCAGAAAGTCAATGCTTGT 2053
Db 1998 AUGAAGGUGUUGAUGCCAGAGAGAGUUAUAUUGACUGGUCUUGUGGAGUACUCCG 2057
Qy 2054 ACGAATGCTAACTTAGAGCACAGGAGTTGGAGTCCCTCAACGATTTCCATAGGCTTGC 2113
Db 2058 GAGUGUCUUAUUCUAAAGAACGAGAGAUAGAGUCUUAAGACGACUUAUAGUCCAAAG 2117
Qy 2114 GTGATAGTGTGATTAACAAGCAATGGCATCGGTGTCTACATGGCTCACTCAAGTT 2173
Db 2118 GCGAUUCUGUAUUGAAGCAGAGUAGUCUGUUAUACACGCGGUCGGAUUAAGUU 2177
Qy 2174 CAACAAATCAAGAACTATGTGACAGTTTGGCAGTCTGCTCGCCACTCTATCAAT 2233
Db 2178 CAGCAAAUAGAAACUUUAUCUAUAGCUGGUAUACUAUUGUCUGCGGUGUGGAU 2237
Qy 2234 CTATGCAAGTCACTAAGGATGAAGTCCGGTATGATTTCTGATTCAGGGAGAAAGTTGT 2293
Db 2238 CUGUCAAGUCCUAAAGAUACAGCUGCUUAGACCUUUAACCGUCAAAGUUGGA 2297
Qy 2294 GTTGGGATGTCATTTGAAAGTGGCTCCTCAACCTGGCGGCAAGGTCAATCATGG 2353
Db 2298 GUCUGGAGUUGCAUCUAGGAAGUGUUAUUCAAACCAACGCGCAAGAGUACUAGG 2357
Qy 2354 GAGATTGCTCTGATTTACAAGGGGAAATGTTTACTGCACTTCTATCTTATCAAGAGAT 2413
Db 2358 GUGUUGUUAACCCACCGAGGAAGUUAUUAUGUGGCGCUUUGGAUAUAGAGCAG 2417
Qy 2414 AGAATGTTGACTGAGAGCGACTGGAGAGGGTGTATCTATCTATCAATGATATAT 2473
Db 2418 GUGUGUGUACAGUGCAUUGAGAGAGAGUAGUGUGUAGCUGCUGAGUCUGUUAU 2477
Qy 2474 TCTGATATTGCAAGCTCAAAATCTGAGGAAACAAATGAGAGAGGTGAACCCACGNA 2533
Db 2478 UCCGCAAGGCGGAAACUCAGAAACUCUGCGCAGACUGUCUGGAAACCGGAGUCC 2537
Qy 2534 CTAATGCAAGATGCTACTTGTGATGGGTGCTGCTGTGTGAAAGTACAAAGAGAT 2593
Db 2538 AGUAGGCAAGAGUUGUUCUGAGCGAGUUCGGGUGUGGAAACCAAGAAUU 2597
Qy 2594 TTTGAAAGATTGATCTTGATGAGATTGATCTTGGTTCTTGGAAACAAAGCTGCTGT 2653
Db 2598 CUUCCAGGUAUUAUUGAAGAGAUUAUUAUAGUACCUUGGAAAGCAAGCCGCGAA 2657
Qy 2654 ATGATCAGAGAGGCTTAATCTGACTGATAAGAGCCACATGCAATGCTGATA 2713
Db 2658 AUGAUCAGAGAGCGUGCGAAUUCUACAGGAUUAUUGUGGCGCACGAGGACACGUAAA 2717
Qy 2714 ACGGTAGATTCACTTCTAATGATCCAAACCGCGATCAC-----ACAAGAGGCTT 2764
Db 2718 ACGUUAUUCUUAUGAUAUUAUUGGAAAGACACACGUGUCUGUUAAGAGGUUA 2777
Qy 2765 TTTATTTGATGAGGTTGATGCTGCACACCGGTTGTTAACTTCTGTTGCTTATCTCT 2824
Db 2778 UUAUUGAAGAGGUGUAGUUGCAUACUGGUGUGUUAUUAUUGUGGCGAUGUCA 2837
Qy 2825 GTTGGCATCGCATACATTTACCGAGATACACAGCATCTCTTCAATTAACAGGTT 2884
Db 2838 UUGUGGAAAUUGCAUUGUUAUCGAGACAACAGCAUUAUCCAUACAUAUAGAGUU 2897
Qy 2885 CAGAATTTCCCGTATCCCAACATTTTGAGAGCTGCAAGTGGATGAGTTGAGATGAGG 2944
Db 2898 UCAGGAUUCUCCGUACCCCGCCCAUUAUUGCCAAUUGGAUUGACGAGGUGGAGACGC 2957
Qy 2945 AGACACACCTGAGATGCCAGGTGATGTAATTTTTCTTCAATCTGAGATGAGGA 3004
Db 2958 AGAAUUAUCUUGUGUUGCCAGCGAUGUACAUUAUUCUGAAACAGGAGAUUAGGGC 3017
Qy 3005 GCGGTGACACCACTTCACTGACACCGATCGGTCTCATCTGAGATGATAGCGGTAG 3064

Db 3018 UUUUGAUGAGCAGUUCUUCGUAUAAAGUUCUUGUUCGAGGAGUUGCGCGGAGCC 3077
Qy 3065 GGAGTACTAAACAGTGTTCCTCAAAACCACTAAAGGGGAAATTTGTAATTTTCACTCAGGCT 3124
Db 3078 GCCGUGAUAUCCGAGUCUCAAACCCUUGCAUGGCAAGAUCCUGACUUAUACCCAAUCG 3137
Qy 3125 GATAAATTTAGTTAGAGGAGAAAGGCTATAGAAATGTGAACACCGTTCTATGAGATCCAA 3184
Db 3138 GAUAAAGAGCUCUGCUUUAAGAGGGUUAUCAGAUUGUUCACUGUGCAUGAAGUCAA 3197
Qy 3185 GGAGAAACCTTTGAAGATGTGCTGCTGATGACGGCAACTCCACTGACTCTGTAT 3244
Db 3198 GCGGAGACAUACUCUGAUGUUUCACUAGUAGGUUAACCCUACACCAAGUUCUUAU 3257
Qy 3245 TCCAAGTCTTCCCGCATGTTCTAGTCTCTGACTAGACACACAAAGAGCTTCAAAAT 3304
Db 3258 CGAGGAGACGCCACAUUUGUGCGAUUGUACAGGCACACCUUUGCGCUCAAGUAC 3317
Qy 3305 TACACCGTAGTGTAGATCCTTTTAGTACAGATAATAGTAGTTGTCTCTTTTAAGCTCC 3364
Db 3318 UACACUGUUAUGGAUCCUUAUAGUUAUUAUAGAGAUCAUAGAGAUCAUAGAGAAACUAGCUG 3377
Qy 3365 TCTCTTTTAGAAATGTATATGCTAGAACAGCTAGTAGCAATTTACAGATGATGCA 3424
Db 3378 UACUUGUAGAUUAGUUAUAGUGCAUGCAGAGAACACAAUAGCAUUAUAGAUAGUACUG 3437
Qy 3425 GTTGTCAAAGGTCAATAATCTCTTTGTGGCAACACCTAAATCAGGAGACTTTCCAGATCTA 3484
Db 3438 GUGUUAAGGUUCCAAUUCUUAUUGUGCAGCGCCAAAGACUGGUGAUUUAUUGAUUG 3497
Qy 3485 CAGTCTATTAGATGTATGCTCCTCGTATAGTACTATCTTAAACAAGTATGATGCT 3544
Db 3498 CAGUUAUUAUGAUUAGUGUCUCCAGGCAACAGCACCAUGAUAUUAUUGAUGUCU 3557
Qy 3545 GTTACCATGAGTTACGTGATAATAGTCTTAATGTGAAGGATGTGTTCTTGATTTTCC 3604
Db 3558 GUUACCAUGAGUUGAUGAUAUUAUUGAUGUCAAUUGAUAUUGAUAUUGU 3617
Qy 3605 AAAAGTATTCGATGTCCAAAGAGGTGAAACCATGTCTAGAGCCAGTTTTCGCTACCGG 3664
Db 3618 AAGUCUGUGCGCCUAGGAUCAAUCAAACCAUUAUACCUUAUUGUACGAAACGCGG 3677
Qy 3665 GCGGAACCGCAAGGCTGACGAGTACTCGAATATCTGTTGCAATGATTAAGAAAT 3724
Db 3678 GCAGAAUUGCCACGCGCAGACUGGACUUAUGGAAAUUAUUGUGCGCAUUAUUAAGGAC 3737
Qy 3725 TTCAACGACACAGACTGACGCGGACGATTGACATTTGAGAGCACCGCATCTCTTTAGTA 3784
Db 3738 UUUAAACGACCCGAGUUGUGUGCAUUAUUGAUAUUGAUAUUGAUAUUGAUAUUG 3797
Qy 3785 GATAAGTTTGTGATGCTATTTTATTAAGAAAGAAATAACAACAAATATTTCTGGA 3844
Db 3798 GAUAAAGUUUUAUUAUUAUUGCUUAAAGAAAGAAAGAAACCAUUAUUAUUAUUG 3857
Qy 3845 GTGATGACGAGGATTTCAATGATGATGTTGGAAACAGAGAAAGATCTATTGAC 3904
Db 3858 UUGUUCAGUAGAGUCUCUCAAUAGUGGUUAGAAAGAGGAAACAGGUAACAAUAGG 3917
Qy 3905 GACTTGGCTAACTACAATTTTACAGATCTCGCGGCATCTGATCAAGACATGATC 3964
Db 3918 CAGCUGCGAGAUUUAUUAUUGAUAUUGCCAGCAGUUGAUAUUAUUAUUAUUAU 3977
Qy 3965 AAGGCTCAACCAAAACAGAAATTTGACCTTTCAATTTCAAGATGAATATCCCTGCTGCA 4024
Db 3978 AAAGCACAACCCAGCAAAAUUUGGACACUUAUUAUUAUUAUUAUUAUUAUUAU 4037
Qy 4025 ACAATTTCTACCATTTCCGAGCAGATCAACGATTTTGGCGGTTCT---CAGAGCTT 4081
Db 4038 ACGAUUGUGUACCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 4097
Qy 4082 ACAAGTTGCTGCTCGAGCATTTGATTTCTTAAGAAAGTTTCTTTTCTTACTAGGAAATCT 4141
Db 4098 ACUAGGCAUAUACUGGACAGUUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 4157

GenCore version 5.1.6
Copyright (c) 1993 - 2005 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: January 17, 2005, 06:20:38 ; Search time 341.911 Seconds
(without alignments)
13211.223 Million cell updates/sec

Title: US-09-551-494-5
Perfect score: 6355
Sequence: 1 gatgttttaatagttttcga.....taaccgcgtagcgccca 6355

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 824507 seqs, 355394441 residues

Total number of hits satisfying chosen parameters: 1649014

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents NA.*

- 1: /cgn2_6/ptodata/1/ina/5A_COMB.seq.*
- 2: /cgn2_6/ptodata/1/ina/5B_COMB.seq.*
- 3: /cgn2_6/ptodata/1/ina/6A_COMB.seq.*
- 4: /cgn2_6/ptodata/1/ina/6B_COMB.seq.*
- 5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq.*
- 6: /cgn2_6/ptodata/1/ina/backfiles1.seq.*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	2468	38.8	6395	3	US-09-259-741-1
2	2468	38.8	6395	3	US-09-037-751-1
3	2468	38.8	6395	3	US-09-466-422-1
4	2468	38.8	6395	4	US-09-962-527-1
5	2466.4	38.8	6395	2	US-08-687-559-2
6	2466.4	38.8	6395	4	US-09-401-415-2
7	2455.4	38.6	6439	3	US-09-259-741-2
8	2455.4	38.6	6439	3	US-09-037-751-2
9	2455.4	38.6	6439	3	US-09-466-422-2
10	2455.4	38.6	6439	4	US-09-962-527-2
11	2452.2	38.6	6475	3	US-09-259-741-4
12	2452.2	38.6	6475	3	US-09-037-751-4
13	2452.2	38.6	6475	3	US-09-466-422-4
14	2452.2	38.6	6475	4	US-09-962-527-4
15	2449.8	38.5	6446	3	US-09-259-741-5
16	2449.8	38.5	6446	3	US-09-037-751-5
17	2449.8	38.5	6446	3	US-09-466-422-5
18	2449.8	38.5	6446	4	US-09-962-527-5
19	2431.2	38.3	6425	3	US-09-259-741-3
20	2431.2	38.3	6425	3	US-09-037-751-3
21	2431.2	38.3	6425	3	US-09-466-422-3
22	2431.2	38.3	6425	4	US-09-962-527-3
23	2286.6	36.0	7685	3	US-09-502-710-22
24	2286.6	36.0	7685	3	US-09-502-710-25
25	2286.6	36.0	7685	3	US-09-502-711-22
26	2286.6	36.0	7685	3	US-09-502-711-25
27	2286.6	36.0	7685	4	US-09-565-616A-1

28 2286.6 36.0 7686 3 US-09-502-710-23
29 2286.6 36.0 7686 3 US-09-502-711-23
30 2286.6 36.0 7687 3 US-09-502-710-24
31 2286.6 36.0 7687 3 US-09-502-711-24
32 2286.6 36.0 7688 3 US-09-502-710-27
33 2286.6 36.0 7688 3 US-09-502-711-27
34 2285 36.0 7926 3 US-09-500-554-1
35 2285 36.0 7926 4 US-09-726-648-1
36 2285 36.0 7926 4 US-10-119-330-1
37 2283.4 35.9 7686 3 US-09-502-710-26
38 2283.4 35.9 7686 4 US-09-502-711-26
39 2283.4 35.9 7686 4 US-09-565-616A-2
40 693.2 10.9 1425 1 US-08-488-672-4
41 690 10.9 1425 4 US-08-788-354-4
42 302.4 4.8 2173 2 US-08-553-619B-6
43 270.2 4.3 1825 1 US-08-176-414B-1
44 270.2 4.3 1825 2 US-08-336-724-1
45 267.8 4.2 807 4 US-09-565-616A-3

ALIGNMENTS

RESULT 1
US-09-259-741-1
; Sequence 1, Application US/09259741
; Patent No. 6033895
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; APPLICANT: HOLTZ, R. BARRY
; APPLICANT: MCCULLOCH, MICHAEL
; APPLICANT: TURPEN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT
; TITLE OF INVENTION: SOURCES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA: US/09/259,741
; APPLICATION NUMBER: US/09/259,741
; FILING DATE: February 25, 1999
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751
; FILING DATE: March 10, 1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
; REFERENCE/DOCKET NUMBER: 00801.0140.US01
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8100
; TELEFAX: 650-463-8400
; TELEX:
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: Genomic RNA
US-09-259-741-1

Query Match 38.8%; Score 2468; DB 3; Length 6395;

QY 2174 CAAATAATGAGAACTATGTGCGACGTTTGGAGCTTCTGTTGTGCGCACATGATCAAAAT 2233
D5 2178 CAGCAAAUAGAAAAACUUUUCGAUAGCUGGUGAGCAUCAUAUUGUGUGCGGUGUCGAAU 2237
QY 2234 CTATGCAAGTCACATAAGGATGAACTCGGTATGATTTCTGATTTCCAGGAGAAAGTTGGT 2293
D5 2238 CUCGUGAAGUCCUCAAAGAAUACAGCUGCUGUAUUGACCUUGAAACCCCGUCAAAGUUGGGA 2297
QY 2294 GTTTGGGATGTCACTTTTCAAAAAGTGGCTCTCAAAACCTGCGGCGCAAAAGTCAATTCATGG 2353
D5 2298 GUCUUGGAUGUUGCAUCUAGGAGUGGUUAUCAAACCAACGCGCAAGAGUCAUGCAUGG 2357
QY 2354 GGAGTTGTCGTGATTAACAAGGGGAAATGTTTACTGCACTTCTATCTTATGAAAGAGAT 2413
D5 2358 GGUGUGUUGAAACCCACGCGAGGAAGUAUCAUGUGGCGCUUUUUGGAAUAUGAGAGCAG 2417
QY 2414 AGAATGTTGATGAGAGCGATGAGGAGGCTGCTGTATCATCTGTATACAAATGGTATAT 2473
D5 2418 GGUGUGUGAGAUGCGAUGAUGGAGAAAGUAGCUGUUGAUCUCUGAGUCUGUUGUUUU 2477
QY 2474 TCTGATATTGCAAAAGCTCAAAATCTGAGGAAACAAATGAGAGACGGTGAACCCACGAA 2533
D5 2478 UCCGACUGGCGAAACUCAGAACTUCUGCGCAGACUGCUUCGAAACGGAGAACCGCAUGUC 2537
QY 2534 CCTACTGCAAAAGATGTTATTTGTGATGGGTGCTGCTGTTGTGGAAGTAAACAAAGGAGAT 2593
D5 2538 AGUAGCGCAAGGUGUUCUUGUGACGAGGUTUCGGGCTUGGGAAGAAACCAAGAAAU 2597
QY 2594 TTTGAAAGATTGATCTTGATGAGGATTTGATCTTGGTCTTGGAAACAAGCTGCTGCT 2653
D5 2598 CUUCCAGGGUUUUAUUUGAUGAAGUAUUAUUUAGUACCTUGGGAAGCAAGCGCGGAA 2657
QY 2654 ATGATCAGAAGAGGGCTAAATCTCTGGAAGTGAATGAGAGCACAAATGGAATGTGAGA 2713
D5 2658 AUGAUCAGAAAGCUGGCGAAUUCUCAGGGAUUAUUGUGGCAACGAGGAAACAGUUNAA 2717
QY 2714 ACGGTAGATTCACTTCTAATGAT-----CCAAACCGCGATCAACAAGAGGCTT 2764
D5 2718 ACCGUGAUCUUCUUGAUGAAUUUUGGGAAGAACACACGUGUCAGUUCAGAGGUUA 2777
QY 2765 TTTATTGATGAGGGTGTATGCTGCACACCGTGTGTAACTTCTGCTGCTTATCTCT 2824
D5 2778 UUCAUUGAAGAGGUGAUGUGCAUCUGGUGUGUUAUUUUCUGGCGCAUGUCA 2837
QY 2825 GGTTCGACATCGCATACATTTACGAGATACACAGACAGATTCTTTTCAATTAACAGAGTT 2884
D5 2838 UUGUGGGAUUUGCAUUGUUUACGAGACACACAGCAGAUUCCAUACAUUAGAGUU 2897
QY 2885 CAGAAATTCCTGATCCCAACATTTTGAAGCTGCAAGTGGATGAAGTTGAGATGAGG 2944
D5 2898 UCAGGAUUCGGAUACCCCGCCCAUUUUGCCAAAUUGGAUUGACGAGGUGGAGACACGC 2957
QY 2945 AGGACCACTGAGATGCGCAGGTGATGTGAATTTTCTTCAATCGAAGTACGAAGGA 3004
D5 2958 AGAACAUCUCCGUGUGUCGCGCAGUGUCAACAUAUUGAAACAGGAGUAUUGAGGCG 3017
QY 3005 GCGGTGACACACACTTCAACTGTATCAACGATCGGTCTCATCTGAGATCATAGGCGGTAAG 3064
D5 3018 UUGUCAUGAGCAUUUUCGUAUAAAGUUCUUGUUCGAGGAGUAGGUGCGCGAGGCC 3077
QY 3065 GGAGTACTAAACAGTGTTTTCAAAACCACTAAAGGGGAAATTTGTAATCTTCACTCAGGCT 3124
D5 3078 GCCGUGAUCUACUUGCAUUAACCCUUGCAUGGCAAGAUCCUGACUUUUAUCCCAUUG 3137
QY 3125 GATAAATTTGATTAGAGAGAGAGGGCTATAGAAATGTGAACACCGTTCTATGAGATCCAA 3184
D5 3138 GAUAAAGAGCUCUCUUUCAAGAGGUAUUCAGAUUGUCAUCUGUGAUGAAGUGCAA 3197
QY 3185 GGAGAAACCTTTGAGATGTGTGCTGTGATGAGGAGCTTCACTGACTCTGATTT 3244
D5 3198 GCGGAGACAUCUCUGAUGUUUUCUAGUAGGUUAACCCUACACACAGUCUCUCCAUCAU 3257

QY 3245 TCCAGTCTTCCCGCGATGTTCTAGTCTGCTGTAGACACACAAAGAGCTTCAAAATAT 3304
D5 3258 GCAGGAGACAGCCCAUUGUUUGGCGCAUUGUCNAAGGCAACACUGUUCUCUCAAUGAC 3317
QY 3305 TACACCGTAGTGTAGATCTTTTAGTACAGATAATTTAGTTGTTCTTCTTTTAAGCTCC 3364
D5 3318 UACACUGUUGUUGAUCUUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3377
QY 3365 TTCTCTTTTGAATAATGTTAGAGAGGTTAGTATAGTATAGCAATTTACAGATGGATGCA 3424
D5 3378 UACUUGUAGAUUAGUUAAGGUGCAGGAAACAUAUAGCAUUAUUAUUAUUAUUAUUAUUA 3437
QY 3425 GTGTTTCAAGGTATATCTTTTGTGCAACCACTAAATCAGAGACTTTTCCAGATCTA 3484
D5 3438 GUGUCAAAGGUUCCAUUUUUUGGCGAGCCAAAGACUGGUGAUAUUUUGAUAUUG 3497
QY 3485 CAGTCTTATTACGATGTATGCTCCCTGGTAAATAGTACTATATCTTAAACAAGTATGATGCT 3544
D5 3498 CAGUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3557
QY 3545 GTTACCATGAGGTTACGTTGATAATAGTCTTAAATGTGAAGGATTTGTTCTTTGATTTTCC 3604
D5 3558 GTUACCAUGAGGUGACUGACAUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3617
QY 3605 AAAAGTATTCCGATGCGCAAAAGGAGGTGAAACCATGTCTAGAGCCAGTTTTCGTACCCGG 3664
D5 3618 AAGUCUGUGCGCGCTTAAAGGAUCAAUAAUAAUAAUAAUAAUAAUAAUAAUAAUAAUAA 3677
QY 3665 GCGGAACCGCCAGGGCTGCAGGACTACTCGAAATCTGTTGCAATGATTAAAGAAT 3724
D5 3678 GCAGAAUUGCCACCCAGACUGGACUAUUGGAAAUUUUAGUGGCGAUGAUAUAAAGAAAC 3737
QY 3725 TTCAACCCACAGACCTGACGGGACGATTGACATTGAGAGCACCGCATCTGTTGTAGTA 3784
D5 3738 UUUAAACGACCGAGUUGUCUGCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3797
QY 3785 GATGAAGTTTTTGTAGTACTTTTATTAATAAAGAAATAACAAAAAATATTTGCTGGA 3844
D5 3798 GAUAAUUUUUUGAUGAUAUUAUUUUUAAAGAAAGAAAGAAAGAAAGAAAGAAAGAAAG 3857
QY 3845 GTGATGACGAGGATTCAATGATGAGATGTTGAAAGAACAGAAAGAGTACTATTGGAC 3904
D5 3858 UUGUUCAGUAGAGAGUCUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3917
QY 3905 GACTTGGCTAACTACAATTTTACAGATCTGCGGCGCATTCGATCAGTACAAGCAATATGC 3964
D5 3918 CAGCUCGAGAUUUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3977
QY 3965 AAGGCTCAACCAAAACAGAAATTTGACCTTTCAATTCAGATGAATACCTCTGCTGCAA 4024
D5 3978 AAAGCACAACCAAAAGUUGGACACUUAUCCAAUCCAAACGAGUACCCCGCUUUGCAG 4037
QY 4025 ACAATTTGTACCAATTCGAAGCAGATCAACGGTATTTTGGCGGTTTCTCA---GAGCTT 4081
D5 4038 ACAGUUGUGUACCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 4097
QY 4082 ACAAGGTTGCTGCTCGAGGCAATTTGATTTCTAAGAAAGTTTCTTTTCTTACAGGAAAT 4141
D5 4098 ACUAGGCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4157
QY 4142 CCAGAAACAGATTTCAAGAAATTTTTCGGATCTCACTCGCAGCTTCTCTATGATGTGTTA 4201
D5 4158 CCAGCGCAGUAGGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4217
QY 4202 GAACTGGATATTTCTAAGTATGATAAGTCAAGAACGAGTTTCAATTTGCTGTAGATAT 4261
D5 4218 GAGCUGGAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 4277
QY 4262 GAAATATGGAAGAGATTTGGGTCTCAATGAGTTTTTGGCGGAGTGTGGAACAAGGGCAC 4321
D5 4278 GAGAUUGCGGAGAGUUGGUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 4337
QY 4322 AGGAAAAACAATTTGAAGGATTAATTGCTGGAATCAAGACATGCTCTGTGTATCAAAGG 4381

Db 798 UUGAGCGAAUCAAACGGCGUGUUUUUUGCGGGAUGGAGACAAGAUUGACCUUUUUUUUGCA 857
Qy 860 GATGAAAGTACTTTAAATATATAGTCATAAATAACAAAATATCTTGCATTTATGTAGTTAAA 919
Db 858 UCAGAGAGUACUUAUAUUGCAUAGUUAUUCUAAUUAUUCUUAAGUUGUGCAAA 917
Qy 920 TCTTACTTTCTGCTTCTAGTAGTAATAGTTTACTTTTAAAGAAATTTTATAGTCACTAGGGTT 979
Db 918 ACUUAUUCUCCGGCCUUAUAGAGAGGUUAUACUAGAAGGAGUUUUUAGUACACAGAGUU 977
Qy 980 AATACTTGGTTTCTTAAATTTACCAAGTAGATACCTATATCTTGTAACAAGAGTTTAGA 1039
Db 978 AAUACCUUGUUUUUGAAGUUUUUUAAGAAUGAUAUUUUUUUUUGUAUAAAGGUGGCGC 1037
Qy 1040 CAAGTAGGGTGTAGTAGTATAGTCTTATGAGCGATGGAAGACGCTTTTGTCTTACAG 1099
Db 1038 CAUAAAAGUGUAGUAGUAGACAGUUUAUACUGCAUUGAAGACGCAUGSCAUUACAA 1097
Qy 1100 AAAACCTTGGCCATGTTCAACACTGAAGAGCAATCTTTAGAGACACGGCTTCGGTTAAC 1159
Db 1098 AAGACUCUUGCAUUGUGCAACAGCGAGAGAAUCCUCCUUGAGGAUUAUACUACAGUCAU 1157
Qy 1160 TTTTGGTTCCCTAAGATGAAGGACATGATGATAGTACCGCTGTTGAGGGTCTATTACC 1219
Db 1158 UACUGUUUUCCAAAAGAGGGAUUGUUAUUGUACCAUUAUUGCAUUAUUUUUGAG 1217
Qy 1220 AGCAAAAAGATGACAAGGAGTGAAGTCAATCTGTAATCGTGACTTCGTTTACACAGTGT 1279
Db 1218 ACUAGUAGAGGACGCGCAAGAGUCUUAUGUCCCAAGGAUUUGUUUAUACAGUCUU 1277
Qy 1280 AATCATATCAGAAATATCAAGCCAAAGCGTTAACTTACAGAAAGTATATCTTTCGTG 1339
Db 1278 AACCACAUUCGAAUACUACAGCGGAAAGCUCUUAUACGCAAAUUGUUUUUGCCUUGUC 1337
Qy 1340 GAGTCTATAGATCCCGTGATATCAATCGTGTACTGCTAGTCTGTAATGCGATGTA 1399
Db 1338 GAUCGAUUCUGAGGGAUUAUUAUACGGUGGACGCGAGGUGCCGAUUGGAUGUG 1397
Qy 1400 GATAAGCAATCTTCAACCCCTGTGATGACTTTCTTCTTGAGACTTAAGCTGGCTGCG 1459
Db 1398 GACAAUUCUUGUUAUAAUCCUUGUCCAGUAGGUUAUACUUGCAUUAUAGUCCGUU 1457
Qy 1460 CTTCAAGACGATATAGTAATCGGAAAGTTTCGGTGTGTTGGATATAGACCACTTCTGAAT 1519
Db 1458 CUAAAGGAUGACUUAUGAUAGCAUUAUAGUCUGUUCGUAUUGGAAACGGUGGCGAGCAU 1517
Qy 1520 ATTTGGGATGAGTGGGCAATTTTGGAAACGTTTCCCACTATCAAGAGAGATTG 1579
Db 1518 GUGUGGAUGAGAUUUUGCGUGGCUUUGGGAAACGAUUUCCUCCGUGAAAGAGAGACUC 1577
Qy 1580 GTGAGCGAGAAATTTCTGATGTAAAGTGAAGTCTCTGAAGATCAAGATCCAGATCTG 1639
Db 1578 UUGAACAGGAACUUAUCAGAGUGGCGGCGACGCAUUAUAGAGUACAGGUGCCUGAUUA 1637
Qy 1640 TATGTCATCGAAAGACGTTTCGTAGTCAATATACCAAGTCTGAGAGTTTACCGCAT 1699
Db 1638 UAUGUGACCUUCCACGACAGAUUAGUACUGAGUACAGGCGCUCUGUGGCAUUGCCUGCG 1697
Qy 1700 CTAGATATCAAGAGGACTTAGAAGAGCTGAGCAAAAGTACGACGCGTTATCAGATTA 1759
Db 1698 CUUGACUUAAGAAAGAGUAGAAACGGAAGUGAUGUAUUAUUGCAUUAUUGCAUUAUUG 1757
Qy 1760 TCTATCTTAAAGGTGCTGATAATTTCCGATATCGCAAGTTTCAAAAGACATGTGCAAGGCT 1819
Db 1758 UCGGUGUUAAGGAGUCUGACAAUUCGUAUGUUAUGUUAUUUUUCCAGAUUGGCCAUUCU 1817
Qy 1820 TTAGATGTAGTCTGTAGTGTGGCAGCACGAGTAATCGTTGAGTGGCGGAGAAATAGAGC 1879
Db 1818 UUGGAAGUUGACCAUAGACGCGAGGAAAGGUUAUAGUCGCGUACUAGAGCAUUGAGAGC 1877
Qy 1880 GGTTAATCTTCTATTTGATAGCCACCGAGAGATGTGGCTTAAGGCTCT-----T 1933
Db 1878 GGUTGACUCUCAUUAUUGAACGACCUUACUGAGCGGAUUGUUGCGUAGCUUAUAGGAG 1937

Qy 1934 AAAAGCAGCGCGTCTGAGCGCGTGTGTATGTCTTGAACCGACATCCGAGAGGGTGAAACGTA 1993
Db 1938 CAAGAGAAGGCUUCAGAGAAGGUGCAUUGGUAGUUAUCCUCAAGAGAAGUUGAAGAACCUGCC 1997
Qy 1994 AATAAATTTTCTATTGCTGAGAAAGGAGATTGCTGTGTGTGCAGAAAGTCAATGGTTTG 2053
Db 1998 AUGAAGGUGUGGAGGCGCAGAGAGAGUUAUUAUUGUGUCUUGUGGAGAUCAUCCG 2057
Qy 2054 ACGAATCTAACTTATAGAGCACGAGAGTTGCGAGTCCCTCAACGATTTTCCATAAGGCTTGC 2113
Db 2058 GAGUGUCUUAUUAUUAAGAACGAGGAGAUAGAGUUAUAGAGAGUUAUUAUUGGCGAGC 2117
Qy 2114 GTGATAGTGTGATTAACAAGCAAAATGGCATCGGTGTGTCTACATCTGGCTCACTCAAAGTT 2173
Db 2118 GCAGUUCGUUAUUCGUAAAGCAGUAGGCGUUAUUAUUAUUAUUAUUAUUAUUAUUAU 2177
Qy 2174 CAACAATGAAAGACTATGTGGACAGTTTGGCAGTTTGGCAGTTCGTTGTCTGCTGCTGATCAAT 2233
Db 2178 CAGAAAUGAAAAACUUUUAUGGAGCCUGGUAUUAUUAUUGUGGUGGUGGUAU 2237
Qy 2234 CTATGCAAGTCACTAAAGGATGAAGTCGGGTATGATTCTGATTCAGGGAGAAAGTTGGT 2293
Db 2238 CUCGUCAGAGUCCUCAAAGUAUACGUGCUUAUUGACCUUGAACCUGCAAAAGUUGGA 2297
Qy 2294 GTTTGGATGTCACTTTTGAAGAAAGTGTCTCTCAAACTCGCGCCAAAGGTCATTCATGG 2353
Db 2298 GUCUUGAUGUUGCAUCUAGAGAGUGGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 2357
Qy 2354 GGAGTGTCTGATTAACAAGGGGAAATGTTTACTGCATCTTCTATCTTATGAAGGAGAT 2413
Db 2358 GGUGUUGUAGAAACCCAGCGAGAGUUAUUAUUGUGGCGUUUUGGAAUUAUAGAGAG 2417
Qy 2414 AGAATGTGATCTGAGAGCGACTGAGAGAGGTGTGTATCATCTGATFACAATGGTATAT 2473
Db 2418 GGUGUGUGACUCCGAGUAGUUGGAGAGAGUAGUAGUAGUAGUAGUAGUAGUAGUAGU 2477
Qy 2474 TCTGATATTCAGAGTCTCAAAATCTGAGGAAACAAATAGAGAGACGTTGAACCCCAAG 2533
Db 2478 UCCGACAUUGCGAAACUCAGAAACUCUGCGCAGACUGUUCGAAACCGAGAACCGCAUGC 2537
Qy 2534 CCTACTCAAGATGTTGATCTTGTGATGGGTGCTGTGTTGTAAGAGTACAAAGGAGAT 2593
Db 2538 AGUAGCCAAAGGUGUUCUUGUGGACGAGUUCGCGGUGUGGAAACCAAAGGAAUUA 2597
Qy 2594 TTTGAAAGATTTGATCTTGTGATGAGATTGATCTTGTGTTCTCGGAAACAAAGCTGCTCT 2653
Db 2598 CUUCCAGGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2657
Qy 2654 ATGATCAGAGAAGGGCTAATTCATCTGGAATGATAAGAGCCACAATGGACAATGTGAGA 2713
Db 2658 AUGAUCAGAGACGUGCGAAUUCUCCAGGGAUUAUUGUGGCCACGAAAGGACACGUAUAA 2717
Qy 2714 ACGTAGATTCATCTTCTAATGCTAT-----CCAAAACCGCATCACACAAGGGCTT 2764
Db 2718 ACCGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 2777
Qy 2765 TTTATGATGAAGGGTGTGCTGCACACCGGTTGTGTAACTTCTCGTGGTCTTATCTCT 2824
Db 2778 UUCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 2837
Qy 2825 GGTTCGACATCGCATATTTACGGAGATACACAGCAGATTCTCTTTCAATTAACAGATT 2884
Db 2838 UUGUGGAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 2897
Qy 2885 CAGATTTCCGCTATCCAAACATTTTGAAGAGTGTGAGGATGAGTTGAGTGAAGG 2944
Db 2898 UCAGAUUCCGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 2957
Qy 2945 AGGACACACTGAGATGCCAGAGTGTGATTTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 3004
Db 2958 AGAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3017

Db 5175 UCUGUGGACAAAAGGAAGGAGCCGACGAGCCACUCUCUGGAUCUUAUACACAGC 5234
Qy 5222 TGCTTGCAAAAGAAATTTCTTTTAAAGCTAATCCCTAATTAATTAATCAATAACATCCGAGGA 5281
Db 5235 AGCUGCAAAAGAAAGAUUUCAGUUCAGGUGUCCCAAUUUGCUUAUAUACCAACCCAGGA 5294
Qy 5282 TGCTGAGNAGCACCGCTGCGAAGTGTAGTGAATTAATCAAGAGTGGCTATGGAAGAGG 5341
Db 5295 CGCAUGAAGAAACGUCUGGCAAGUUAUUAUUAUUAAGAAUUGGAGAGUUGUAGUGUGCAGCGG 5354
Qy 5342 ATACTGTCCTTTATCTTTGCGAGTTCGTTTCAATTTGCTAGTACATATAAAATAATGTAAAG 5401
Db 5355 UUUCUGCCGCUUUCUGGAGUUGUGUGUGUGUUAUUGUUAUUAAGAAUUAUUAUAA 5414
Qy 5402 AAAAGGTTGAGGAAACGTATTTTGAAGTGTACAGACGCGCTCGCAATTTGAACCTACTGA 5461
Db 5415 AUUAGGUUUGAGAGAGAGAUUAUCAAACGUGAGAGCGGAGGCCCAUGGAACUUAACAGA 5474
Qy 5462 AAAGGTTGTTGAGGAGTTCGTTGATGAGTACCAATGCTGTGAACCTCGAAAGGTTTC- 5520
Db 5475 AGAAGUCUGUAGAGUUAUGGAAGUUGUCCUUAUGUGCAUGAGGCUUGCAAAAGUUCG 5534
Qy 5521 -----CGGAAACAAAAAGAAATGCTAGGTAATTAATTTAATAAT----- 5561
Db 5535 AUCUGAACCGGAAAAAGAGUGUUGCCGCAAGGGGAAAAUAGUAGUAGUAGUGGUC 5594
Qy 5562 -----AAGAAAAATAACAGCTGTGTAAGAGGTTTTTAAATTTGAGGAAATTTGAGGA 5614
Db 5595 AGUGCCGAAACAAGAAUAGUUAAGUUAUUGGAGGAGUUAUUGGAGUUAUUAAGAA 5654
Qy 5615 TAATGTAAGTATGACAGT-----CTATCGCTCATCGAGTAGCTTTTATCAAT 5665
Db 5655 UAAUUAUUCGAUGAUGAUGCGAGGCUACUGUGCGCAUUGCGGAUUGUUAUUAUUA 5711
Qy 5666 ATGCTTATACATCAACTCTCGAGCAATTTGTTTACTTATCTTCGCTTACGCGAGT 5725
Db 5712 AUGUUAUACUAUACUAUCCAUUCUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 5771
Qy 5726 CTTGTGACGTGATCAATCTGTGTACAAATGATTCGTTGTAACCGATTTTCAAAACGCAACAA 5785
Db 5772 CCAUAGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 5831
Qy 5786 GCTAGGACACAGTCCAAACAGCAATTTGCGGATGCGTGGAAACCTGTGCTAGTATGACA 5845
Db 5832 GCUGAACUGUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 5891
Qy 5846 GTGAGATTTCTGCACTCGATTTCTATGTGTATAGATATAATTCGACCTTTGATCCGTTG 5905
Db 5892 GUUAGGUUCCUGACAGUGACUUAUAGGUGUACAGGUACAUGCGGUUAUAGACCCGCUA 5951
Qy 5906 ATCAGCGGTTTATAATAGCTTTGATACCTAGAAATAGAAATAATAGAGTTGATATCAA 5965
Db 5952 GUCACAGCACUUGUAGGUGCAUUGACACUAGAAUUAUUAUUAUUAUUAUUAUUAUUA 6011
Qy 5966 CCGCAGCGAATACCTAGTAATCGTTAACCGGACTAGAGGTTAGAGTGTCTACTGTA 6025
Db 6012 GCGAACCCACGACGUGCGGACGCUUAGUAGUACUUGUUGUUGUUGUUGUUGUUGUUG 6071
Qy 6026 GCTATAAGGCTTCAATCAATAATTTGCTTAATGAACTGTTGCTGGAATGCGCATGTC 6085
Db 6072 GCUAAGAGCGGCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 6131
Qy 6086 AATCAAGCAGGCTTTGAGACTGCTAGTGGACTTGTCTGGACCAACACTCCGGCTACTTAG 6145
Db 6132 AAUCGGAGCUUUCUGAGAGCUUUCUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 6188
Qy 6146 CTATTGTTGAGATTCTCTAAATAAGTCTGCTGAAGACTTAAATTTACGGTGGCTGA 6205
Db 6189 -UGAGGUGAUGCAAGUAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 6246
Qy 6206 TACCAAAATCAGAGTGTGTTGCTCCACTTTAATAATAAGATTTGTATATCTGATCC 6265
Db 6247 UACGAUAACCGUAUGUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUGUUG 6305

Qy 6366 AACAGTTAAACCATGTGATGTGTATCTGTGTTATCTGCGTAAACATCGGAGAGTTTCG 6325
Db 6306 CGCGGUCAAAUGUAUUGGUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 6365
Qy 6326 AATCCTCCCTAAACCGCGGTAGCGGCCA 6355
Db 6366 AUCCCCCGUUUACCCCGUAGGGGCCA 6395
RESULT 4
US-09-527-1
; Sequence 1, Application US/09962527
; Patent No. 6740740
; GENERAL INFORMATION:
; APPLICANT: GARGER, STEPHEN
; HOLTZ, R. BARRY
; MCCULLOCH, MICHAEL
; TUREN, THOMAS
; TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
; PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
; FROM PLANT SOURCES
; NUMBER OF SEQUENCES: 5
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Howrey & Simon
; STREET: 1299 Pennsylvania Avenue N.W.
; CITY: Washington
; STATE: DC
; COUNTRY: USA
; ZIP: 20004
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSeq for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/962,527
; FILING DATE: 24-Sep-2001
; CLASSIFICATION: <Unknown>
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 09/037,751
; FILING DATE: 10-march-1998
; ATTORNEY/AGENT INFORMATION:
; NAME: Halluin, Albert P
; REGISTRATION NUMBER: 25,277
; REFERENCE/DOCKET NUMBER: 00801.0140.999
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 650-463-8109
; TELEFAX: 650-463-8400
; TELEX: <Unknown>
; INFORMATION FOR SEQ ID NO: 1:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 6395 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: single
; TOPOLOGY: unknown
; MOLECULE TYPE: Genomic RNA
; SEQUENCE DESCRIPTION: SEQ ID NO: 1:
US-09-527-1

Query Match 38.8%; Score 2468; DB 4; Length 6395;
Best Local Similarity 44.6%; Pred. No. 0;
Matches 2851; Conservative 1173; Mismatches 2300; Indels 66; Gaps 11;
Qy 20 ACAACAACAAATTAACAAACAAACAAACAAATTTATCAACAAACAAACAAACAAATGCGCAC 79
Db 18 ACCAACACAAACAAACAAACAAACAAACAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 77
Qy 80 ATACAATCTATAATTAGCAACGCCCTTTCTTGAAGCGTGTAGTGTGTAACAACTCTCGTT 139
Db 78 ACACAGACAGCUACCAUACUAGCUGUUGUGGACACUUGCCGAGGAAACAAUCCUUGGUC 137
Qy 140 AATGACCTTGCAAGAGCGCATGTACGATACGCCCGCTGGAGAAATTTAAACGCCCGCAC 199

Db 1698 CTTGACATTAGAGAGATCGAAGAACGAGTGAATGATCAATGCACATTTTCAGAGTTA 1757
Qy 1760 TCTATCCCTTAAGGGTCTGTATTAATTTTCATATCCGGAAGTTCAAGACATGTGCAAGCT 1819
Db 1758 TCGGTGTTAAGGGAGTCTGACAAATTCGATGTTGATGTTTTTCCAGATGTGCCAATCT 1817
Qy 1820 TTAGATGTTAGTCTGATGTCGACGACGAGTAAATCGTTGAGTGGCGGCGAAGTAGAAGC 1879
Db 1818 TTGAAGTTGACCCATGACGCGACGAGGTTATAGTCGGGTCATGAGCAATGAGAGC 1877
Qy 1880 GGTTTAACTCTTACTTTTATAAGCCAAACGAGGAGAAATGTGGCTAAAGCTCT-----T 1933
Db 1878 GGTCTGCTCTCACATTTTGAACGACCTACTGAGGCGAATGTGCGCTAGCTTTTACAGGAT 1937
Qy 1934 AAAAGCAGCGGTCTGAGCGCGTGTATGTTTGAACCGCATCCGAAGAGGTGAACGTA 1993
Db 1938 CAAGAGAGGGCTTCAGAAAGGTCTTTGGTAGTTACCTCAAGAGAAATTTGAAGAACCGTCC 1997
Qy 1994 AATAAATTTTCTATTGCTGAGAAAGGAGATTGCCCTGTGTGTCAGAAAGTCATGGTTTG 2053
Db 1998 ATGAAGGTTTCATGCGCCAGAGGAGATTACATTTAGCTGTCTTGTGGAGATCATCCG 2057
Qy 2054 ACGAATGCTTAATTAGACACGAGGAGTTGGAGTCCCTCAACGATTTCCATAAGGCTTGC 2113
Db 2058 GAGTCGTCTTATCTAAGAACGAGGAGATAGAGTCTTTAGAGCAGTTTCATATGGCAACG 2117
Qy 2114 GTGGATAGTGAATTAAGCAAAATGCGTGTGTACACTGGCTCACTCAAGCTT 2173
Db 2118 CGAGATTCGTTAAATTCGTAAGCAGATGAGTCGATTTGTGTACACGGGTCCGATTAAGTT 2177
Qy 2174 CAACAAATGAAGAACTATGTGACAGATTGGCAGCTTGTGTCGCGCACTGTATCAAT 2233
Db 2178 CAGCAAAATGAAAACTTTATCGATAGCTGTGTAGCATCACTATCTGCTCGGCTGCGAAT 2237
Qy 2234 CTATCGAAGTCACTAAGAGTGAAGTCTGGGTATGATTTCTGATTCAGGGAGAAAGTTGGT 2293
Db 2238 CTCGTCAAGATCCTCAAGATACAGCTGCTATTGACCTTTGAACCCCGTCAAAAGTTTGA 2297
Qy 2294 GTTTCGGATGTCATTTGAAAGTGGCTCTCAACCTCGCGCCAAAGGTCATTCATGG 2353
Db 2298 GTCTGGATGTTGCATCTAGGAAGTGGTTAATCAACCAACGCGCCAGAGTCACTGATGG 2357
Qy 2354 GGAGTTGTCCTGGATTTACAAAGGGGAAATGTTTACTGCACTTCTATCTTATGAAGAGAT 2413
Db 2358 GGTGTTGTTGAACCCAGCCGAGGAGTATCATGTGGCGCTTTTGAATATGATGAGCAG 2417
Qy 2414 AGAATGCTGACGAGAGCAGTGGAGGAGGTGGCTGTATCATCTGATACAAATGTTAT 2473
Db 2418 GGTGTGTTGACATGCGATGATTGGAGAGAGTAGCTGTGAGCTCTGCTGTTGTTAT 2477
Qy 2474 TCTGATATTGCAAGCTCCAAATCTGAGGAAACAAATGAGAGACGTTGAACCCACGAA 2533
Db 2478 TCCGATGCGGAAACTCAGAACTCTGCCAGACTGCTCGAAGCGGAGAACCGATGTC 2537
Qy 2534 CCTACTGCAAGATGTTACTTGTGATGGGGTCCCTGGTTGTGGAAGTACAAAGAGAT 2593
Db 2538 AGTAGCGCAAGGTTGTTCTTGTGGACGAGTTCCGGGCTGTGGGAAACCAAGAAAT 2597
Qy 2594 TTTGAAAGATTGATCTTCAATGAGGATTTGATCTTGGTTCTTGGAAACCAAGCTGCT 2653
Db 2598 CTTTCCAGGGTTAAATTTGATGAAGATCTAAATTTTAGTACCTGGGAGCAGCCGCGAA 2657
Qy 2654 ATGATCAGAAAGAGGCTTAATCTGACTGATGAAGAGCCAAATGGAATGCAATGTGAGA 2713
Db 2658 ATGATCAGAAAGAGCTCGAATCTCTCAGGATTTATTTGTGCCACGAGGACCACTTAA 2717
Qy 2714 ACGGTAGATTCACTTTAATGCAAT-----CCAAACCGGATTCACAGAGGCTT 2764
Db 2718 ACCGTTGATCTTTCATGATGAATTTTGGGAAAGCACACGCTGTGAGTTCAAGAGTTA 2777
Qy 2765 TTTATTTGATGAGGTTGATGTCACACCGGTTGTGTTAACTTCTCGTGTCTTATCTCT 2824

Db 2778 TTCATTGATGAAGGGTTGATGTTGCATACCTGTTGTTGTTAAATTTTCTTGTGGGATGTCA 2837
Qy 2825 GGTTCGACATCGCATACATTTTACGAGATACACAGCAGAGATTCTTTTCAITAAACAGAGTT 2884
Db 2838 TTGTGCGAATTTGATATGTTTACGAGACACACAGCAGAGATTCCATACATCAATAGAGTT 2897
Qy 2885 CAGAAATTTCCGTTATCCCAAACTTTTGAAGCTGCAAGTGGATGAAGTTGAGATGAGG 2944
Db 2898 TCAGGATTTCCCGTACCCCGCCCATTTTGCCTTAAAGTTTGAAGTTGACGAGGTGGACACGC 2957
Qy 2945 AGGACACATGAGATCCCGAGGTGATGTAATTTTCTTACAAATCGAAGTACGAAGGA 3004
Db 2958 AGAATCTACTCTCGTTGTCAGCGGATGTACATTTATCTGAACAGAGATATGAGGCG 3017
Qy 3005 GCGGTGCAACCACTTCAACTGTACAACTGTAACGATCGGTCTCATCTGAGATGATAGCGGTAAG 3064
Db 3018 TTTGTCTATGACACTTCTCGGTTAAAGTCTGTTTCGCGAGGAGATGGTCGCGGAGCC 3077
Qy 3065 GGAGTACTAAACAGTGTTCCAAACCACTTAAAGGGAATTTGTAATTTTCACTTCACTCAGGCT 3124
Db 3078 GCCGTGATCAATCCGATCTCAAAACCCCTTGCAATGGCAAGATCCCTGACTTTTACCCAATCG 3137
Qy 3125 GATAAATTTGAGTTAGAGGAGAGGGCTTATAAGAAATGTGAACACCGTTTCATGAGATCCAA 3184
Db 3138 GATAAGAGAGCTCTGCTTTTCAAGAGGTTTTCAGATGTTTACACTGTGCAATGAAGTCAA 3197
Qy 3185 GGAGAAACCTTTGAAGATGTTGCTGCTGATGAGTTGACGGCAACTCCACTGACTCTGATTT 3244
Db 3198 GCGCAGACATACTCTGATGTTTCACTAGTTAGGTTAAACCCCTACACCACTCTCCATCAT 3257
Qy 3245 TCCAAGTCTTCCCGCATGTTCTAGTCTGCTGCTGATGAGTTGACGGCAACTCCACTGACTCTGATTT 3244
Db 3258 GCAGGAGACGCCACCATGTTTGGTCGCTGATGTCAGAGGACACCTGTTCCGCTCAAGTAC 3317
Qy 3305 TACACCGTAGTGTGATGATCCCTTTAGTACAGATAATTTAGTGAATTTGTCTTTTAAAGCTCC 3364
Db 3318 TACACTGTTGATGAGTCTTCTTGTAGTATCATTAGAGATCTAGAGAACTTAGCTCG 3377
Qy 3365 TTCTTTTGAAGATGATATGTTGATGAGAGAGTGTAGTATGATGAATTTAGTGAATTTTAAAGCTCC 3424
Db 3378 TACTGTTAGATATGATATAGGTCGATGACGAGAACCAATAGCAATTTACGATTTGACTCG 3437
Qy 3425 GTGTTCAAGGTCATATCTCTTTGTGGCAACACTTAATCAGGAGACTTTCCAGATCTA 3484
Db 3438 GTGTTCAAGGTTCCAACTCTTTTGTGCGCGCCAAAGACTGTTGATATTTCTGATATG 3497
Qy 3485 CAGTTCTATTACGATGATGCTCCCTGTTAATAGTACTATATCTTAAACAGTATGATGCT 3544
Db 3498 CAGTTTCTATGATTAAGTGTCTCCAGGCAACAGCACCACATGATGAATAATTTTGTATGCT 3557
Qy 3545 GTTACCATGAGGTACGTTGATTAATGATCTTAAATGTAAGGATTTGTTCTTGTGATTTTCC 3604
Db 3558 GTTACCATGAGGTGACTGACATTTTCAATGAAATGTTCAAGATTTGATTTGATATGTTCT 3617
Qy 3605 AAAAGTATTCGATGCCAAAGAGGTGAACCATGTCTAGAGCCAGTTTGTGTTACCCGCG 3664
Db 3618 AAGTCTGTTGTCGCCCTTAAGATCAATCAACCACTTAATACCTATGTTGACAGGCGG 3677
Qy 3665 GCGGAAACCGCCAAAGGGCTGCGAGGACTCTCGAAAACTCTGTTGCAATGATTTAAAGAAAT 3724
Db 3678 GCAGAAATGCCACCGCAGACTGGACTATTGGAAAAATTTAGTGGCGATGATTTAAAGGAAC 3737
Qy 3725 TTCAAACGACAGACTGACGGGAGGATTTGACATTTGAGAGCAGCAGCTCTGTTGATGTA 3784
Db 3738 TTTAAACGACCGGAGTTGTTGCGCATCATTTGATATTGAAAACTGTCATCTTTAGTTGTA 3797
Qy 3785 GATAAGTTTGTGATGCTTATTTTATTAAGAAAGAAAAATAACAATAAATAATTTCTGGA 3844
Db 3798 GATAAGTTTGTGATGATTTTGTCTTAAAGAAABAAAGAAACCAATAAATAATTTTCT 3857
Qy 3845 GTGATGACAGAGGATTTCAATGATGATGTTGGAAAAACAGGAAAGAGTACTATTTGAC 3904
Db 3858 TTGTTTCACTAGAGAGTCTCTCAATAGATGTTTGAAGAAACGAGGAAACAGGTAACATAGGC 3917

Qy	3905	GAC	TTGGCTAAC	TACAAATTTT	TACAGATCTGCGGCCA	TCGATCAGTACAAAGCACATGATC	3964
Db	3918	CAG	CTCGCAGATTTT	TGATTTTGT	TAGATTTGCGCAGCTTGATCAGTACAGACACATGATT	3977	
Qy	3965	AAG	GCTCAACCCAA	ACAGAAATTTG	GCCTTTCAATTCAGATGAATACCCCTGCTCTGCAA	4024	
Db	3978	AA	GCAACACCCAG	CAAAATTTG	GACATTTTCAATCCAAACGGAGTACCCGGCTTTTCGAG	4037	
Qy	4025	ACA	TTGCTCTACCA	TTCCGAAGCAGATCAA	CGGTATTTTGGCCGGTTTCTCA---GAGCTT	4081	
Db	4038	ACG	ATTGTTACCA	TTCAAAAAGATCAATG	CAATATTTGGCCCGTTGTTTAGTGAGCTT	4097	
Qy	4082	ACA	AGTTGCTGCTCG	AGGCATTTGATTTCTA	GAAAGTTCTCTTTTCTTTTACTAGGAAACT	4141	
Db	4098	ACT	AGGCAAATTTACT	GGCAGTGTGATTCG	AGCAGATTTTTTGTGTTTTTCACAAGAAAGACA	4157	
Qy	4142	CC	GAAACAGATTTC	CAAGAAATTTTTCTCG	AGTCTCGACCTCGACCTCGCAGGTTCCTATGSGATGCTGTA	4201	
Db	4158	CC	AGCAGATTTAG	AGATTTCTTCGG	AGATCTCGACAGTCAATGTCGCAGTGAATGCTCTTG	4217	
Qy	4202	GAA	CTGGATATTTCT	TAAAGTAGA	TAAGTCAAGAAACGAGTTTCAATGTGCTGTAGAGTAT	4261	
Db	4218	GAG	CTGGATATATCAA	AATACGACAAATCTC	CAGAAATGAAATCCACTGTGCAGTAGAATAC	4277	
Qy	4262	GAA	ATATGGA	AAAGATTGGGTCTCA	ATGAGTTTTTGGCCGAAGTGTGGAACAAAGGGCAC	4321	
Db	4278	GAG	ATCTGGCG	AGATTTGGGTTTTGA	AGACTTCTTTGGGAGAGTTTGTGGAACAAAGGGCAT	4337	
Qy	4322	AG	AAAAACA	ACTTTGAAGGATTA	CAATTTCTGTGAATCAAGACATGCTGTGCGTATCAAAGG	4381	
Db	4338	AGA	AGACCA	CCCTCAGAGATTA	TACCGCAGGTATAAAACTTGCACTCTGGTATCAAAGA	4397	
Qy	4382	AAA	AGCGGTGATGTG	ACTATTTCA	TCCGCAATACTGTTATAATPAGCAGCTTGCTTGGGT	4441	
Db	4398	AAG	AGCGGGACGTCA	CGAGTTTCA	TGTGGAACACACTGTGTATTTGCTGCAATGTTTGGCC	4457	
Qy	4442	TC	ATGTTACCGATG	GGAAGGTCATAAA	AGGTGCTTTTTTGTGGAAGCAGTACCGTTTTCG	4501	
Db	4458	TC	GATGTTCCGAT	TGGAGAAATAATCAA	AGGAGCCCTTTTTCGGGTGACGATAGTCTGCTG	4517	
Qy	4502	TAT	TTTCCAAAGGGT	TTTGGATTTCC	CTCGACATTCAGTGTGCTAATCTCATGTGGAAT	4561	
Db	4518	TAC	TTTCCAAAGGGT	TGTGAGTTTCC	CGATGTGCAACACTCCGGGATCTTATGTGGAAT	4577	
Qy	4562	TTT	GAGGCCAACTGT	ATAGAAAGGTA	CGGTTACTTTTGTGGTAGATACATCATAC	4621	
Db	4578	TTT	GAAACAAAC	TGTTTAA	AAAAACAGTATGGAATACCTTTTGGGAAAGATATGTAATACAT	4637	
Qy	4622	CAT	GATAAGGG	GCAATAGTGTATATG	ATGATCTTTGAAGTTGATCTCCAACCTTGGGCA	4681	
Db	4638	CAC	ACAGAGAGTGC	ATTTGTGTATTA	CGATCCCTTAAAGTTGATCTCGAACTTGGTGT	4697	
Qy	4682	AAA	CATATCAAG	GATTATGATCA	CTTAGAAGAGTTAAGGGTGTCTTTTGTGCGATGTTGCT	4741	
Db	4698	AAA	CACATCAAG	GATTGGA	ACACTTCGAGGAGTTTCAGAAAGTCTCTTTGTGATGTTGCT	4757	
Qy	4742	TGT	TTCGCTCG	GAACCTG	GTCTAGGCTTTTCCGAGCTGAAACGAGCTATCAAGAGGTT	4801	
Db	4758	GTT	TCGTT--	--GAA	CAATGTGCGGTATTA	CACACAGTTGGACCGCTGTATGGGAGGTT	4814
Qy	4802	CATA	AAACCGCG	ATGATGTTG	CTGTTGCTTTTAAATGCTTAACAAATTTTGTGTGAT	4861	
Db	4815	CAT	AAGACCG	CCCCCTCC	AGGTTGCTTTGTTTATAAAAGTCTGGTGAAGTATTTGTCTGAT	4874	
Qy	4862	AA	ATTTTTTAT	TTTAGA	ACTTTGTTTTTAAATGGCTGTTTAGTCTCAGAGATACTGTCAAAAT	4921	
Db	4875	AA	GTCTCTTTT	TAGA	AGTTTGTTTATAGATGGCTCTAGTTGTTTAAAGAAAGTGAATAT	4934	
Qy	4922	TAG	CAGGTTTC	ATTGATCTTT	TCGAAACAGGATGAGATCTTCCGGCATTCATGACTAAGT	4981	
Db	4935	CA	ATGAGTTTAT	TCGACTCG	CAAAAATGGAAGATCTTACCGTCTGATGTTTACCCCTGT	4994	

Qy	4982	CAAGAGTGTAGAAATATCGACTCTGGACAAAGATTATGGCTGTGTTAAGAAATGATAGTCTTTC	5040
Db	4995	AAAGAGTGTATATGTGTTCCAAAGTTGATAAAAAATAATGGTTCATGAGAAAGAGTCATGTGC	5054
Qy	5042	TGATGTAGATTTACTTTAAAGTGTAAAGTGTAGTTTAAAGAAAGGATATGCTGTGCTTAGCTGTA	5101
Db	5055	AGAGTGAACCTTCTTTAAAGGAGTTAAGCTTATTTGATAGTGGATACGTCTGTTTAGCCGG	5114
Qy	5102	TTTGGTAGTGTCTGGGGAGTGGAAATCTCCCGGATAAATCGCCGTGGTGGTGTCTCAGTGTGTTG	5161
Db	5115	TTTGGTGTCTACGGCGGAGTGGAACTTGCCTGACAAATGCAGAGGAGGTGTGAGCGGTGTG	5174
Qy	5162	TATTTGTAGATAAGAGAAATGAAAAAGAGTAGTAAGGAAGCAACGCTGGGTGCGTATCACGCCCC	5221
Db	5175	TCTGGTGGACAAAAAGGATGAAAAGAGCCGACGAGGGCCACTCTCGGATCTTTACTATCACAGC	5234
Qy	5222	TGCTTGCAAAAAGAAATTTTCTTTTAAAGCTAAATCCCTAAATATTTCATTAACATCCAGGA	5281
Db	5235	AGCTGCAAGAAAGATTTTCAGTTCAAGGTGCTTCCCAATTTATGCTATTAACCAACCCAGGA	5294
Qy	5282	TGCTGAGAACACCCGTGGCAAGTGTTAGTGAATATCAAAGGAGTGGCTATGGAAGAAAGG	5341
Db	5295	CGCGATGAAAAAGTCTGGCAAGTTTGTAGTTAATAATTAGAAATGTAAGATGTCACGGGG	5354
Qy	5342	ATACTGTCTTTATCTTTGGAGTGTGTTTCAAATTTGTGTAGTACATAAAAAATATGTAAG	5401
Db	5355	TTTCTGTCCGCTTCTCTGGAGTTTGTGTCCGTGTGTATTTGTTTATAGAAAATAATATAA	5414
Qy	5402	AAAGGTTTTCAGGGACGTAATTTTGATGTGACAGACGGCTCGCCAAATTTGAACCTACTGA	5461
Db	5415	ATTAGGTTTCAGAGAGAAGATTAACAAAGTGAGACGGAGGGCCCCATCGAACTTACAGA	5474
Qy	5462	AAAGGTTTTCAGGAGTTTCGTGATGAAGTACCAAATGGCTGTGAAACTCGAAAGGTTTC	5520
Db	5475	AGAGTGTGTATGAGTTTCATGGAAGATGTCCCTATGTGATCGATCAGGCTTGCAAGTTTCG	5534
Qy	5521	-----CGGAAAAACAAAAAGAAATGGTAGTAAATAATGTTAAATAAT-----	5561
Db	5535	ATCTCGAAACCGGAAAAAAGAGTGTATGTCGCAAGGGGAAAAATAGTAGTAATGATCGGTC	5594
Qy	5562	-----AAGAAATAATAACAGTGGTAGAGGGGTTTTAAATTTGAGGAAATTCAGGA	5614
Db	5595	AGTGCCGAAACAGAACTATAGAAATGTTAAGGATTTTGGAGGAATGAGTTTTTAAAGAAGAA	5654
Qy	5615	TAATGTAAGTGATGACGAGT-----CTATCGCGGTCAATCGAGTACGTTTTTAAATCAAT	5665
Db	5655	TAATTTAATCGATGATGATTCGAGAGCTACTGTGCGCGAATCGGATTCGTTTTA---AAT	5711
Qy	5666	ATGCTTTATACAATCAACTCTCCGAGCCAAATTTTGTACTTATCTTCCTCCGCTTACGAGAT	5725
Db	5712	ATGTCTTTACAGTATCACTACTCCATCTCAGTTCGTGTCTTTGTCTATCAGCGTGGGCCGAC	5771
Qy	5726	CCTGTGACAGCTGATCAATCTGTGTACMAATGCAATTCGGGTAAACGAGTTTCAACGCAACAA	5785
Db	5772	CCAATAGAGTTAATTAATTTATGTACTTAATGCTTATGCTTAGGAAATCAGTTTTCAACACAA	5831
Qy	5786	GCTAGGACAACTGTCAAACAGCAATTTGCGGATGCTCGAAACCTGTGCTCTAGTATGACA	5845
Db	5832	GCTCGAACTCTGTTTCAAGACAAATTCAGTGAGGTGTGGAACTTCCACCAAGTAACT	5891
Qy	5846	GTGAGATTTCTGCATCGGATTTCTATGTGTATAGATATAATTCGACGCTTGATCCGTTG	5905
Db	5892	GTTAGGTTTCCCTGACAGTGACTTTTAAAGGTGTACAGGTACAAATGCGGTATTTAGACCCGCTA	5951
Qy	5906	ATACGCGGTTATTAATAGCTTTTGATCTAGAAATAGAAATAATAGAGGTTGATATATCAA	5965
Db	5952	GTCAACAGCACTGTTAGTGCAATTCGACACTAGAAATAGAAATAATAGAAAGTTGAAATACAG	6011
Qy	5966	CCGCAACCGAATTAATCTGAAATCGTTTAAACGCGACTCAGAGGGTAGACGATGCTACTGTA	6025
Db	6012	CGGAACCCACAGACTGCCGAAACGTTAGATGCTACTCTGTAAGTAGACGACGCAACGGTG	6071
Qy	6026	GCTATAGGGCTTCAATCAATAATTTTGGCTTAATGAACCTGGTTCGTGGAACTGCGCATGTTTC	6085

Db 3198 GCGGAGACATCTCTGATGTTTCACTAGTTAGGTTAAACCCCTACACAGCTCTCCATCAT 3257
Qy 3245 TCCAAGTCTTCCCGCATGTTCTAGTCTGCTCTGACTAGACACACAAAGAGCTTCAATAT 3304
Db 3258 GCAGGAGACAGCCACATGTTTGGTGCATTTGTCAAGGCACACCTGTTCTCAAGTAC 3317
Qy 3305 TACACCGTAGTGTAGTACCTTTAGTACAGATAATTTAGTGAATTTGCTCTTTAAAGCTCC 3364
Db 3318 TACACTGTGTATAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3377
Qy 3365 TTCTCTTTAGAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3424
Db 3378 TACTGTTAGATATGATAAGGTCGATGATGATGATGATGATGATGATGATGATGATGATG 3437
Qy 3425 GTGTTCAAAGGTCAATCTCTTTGTGGCAACACCTAAATCAGGAGACTTTCCAGATCTA 3484
Db 3438 GTGTTCAAAGGTCCAATCTTTTGTGCGCCAAAGATGATGATGATGATGATGATGATG 3497
Qy 3485 CAGTTCTATACGATGATGCTCCTCGTGAATAGTATGATGATGATGATGATGATGATG 3544
Db 3498 CAGTTTCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3557
Qy 3545 GTTACCATGAGGTTACGTGATATAGTCTTAATGTGAAGGATGATGATGATGATGATG 3604
Db 3558 GTTACCATGAGGTTGACTGATGATGATGATGATGATGATGATGATGATGATGATG 3617
Qy 3605 AAGATATTCGATGCCAAGAGGTGAACCTGCTAGAGCCAGTCTTGGTACCGCG 3664
Db 3618 AAGTCTGTTGCTGCGCTCAAGGATCAAAATCAACCACTAATCCTATGATGATGATG 3677
Qy 3665 GGGGAACCGCAAGGCTGCGAGTACTCGAATACTGTTGCAATGATTAAGAAAT 3724
Db 3678 GCAGAAATGCCAGCCAGACTGACTATGGAATTTAGTGGCGATGATTAAGGAAC 3737
Qy 3725 TTCAACGCAACAGACTGACGAGGAGATGATGATGATGATGATGATGATGATGATG 3784
Db 3738 TTAAACGCAACCGAGTGTCTGCGCATCAATGATGATGATGATGATGATGATGATG 3797
Qy 3785 GATAGTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3844
Db 3798 GATAGTCTTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3857
Qy 3845 GTGATGACGAGGATTCATGATGATGATGATGATGATGATGATGATGATGATGATG 3904
Db 3858 TTGTTCACTAGAGTCTCTCAATAGATGTTTGAAGAAAGCAGGACAGTAACTAGC 3917
Qy 3905 GACTGGCTAACTACAAATTTTACAGATCTGCGCGCATGATGATGATGATGATGATG 3964
Db 3918 CAGCTCGCAGATTTGATTTTGTAGATTTGCCAGCAGTTGATCAGTACAGACATGAT 3977
Qy 3965 RAGGCTCAACCAAAACAGAAATTTGACCTTTCAATTCAGATGATGATGATGATGATG 4024
Db 3978 AAGACACACCAAGCAAAATTTGACACTTCAATCCAAACCGAGTACCCGCTTTGAG 4037
Qy 4025 ACAATTTGCTACCATTCGAAGCAGATCAACGGTATTTTGGCCGGTTTCTCA --- GAGCTT 4081
Db 4038 ACGATTTGCTACCATTCAAAAAAGATCAATGCAATATTTGGCCGGTTTGTAGTCA 4097
Qy 4082 ACAGGTTCTGCTCGAGCATTTGATTTCTAGAGTCTTTCTTTCTTTCTTTCTTTCTTT 4141
Db 4098 ACTAGGCAATTAAGTACAGTGTGATTCGAGCAGATTTTGTGTTTCAAGAAAGACA 4157
Qy 4142 CAGAACAGATTCAGAAATTTTCTCGGATCTCGACTCGCAGCTCTCTATGATGATGAT 4201
Db 4158 CAGGCGCAGATTCAGAGATTTCTTCGGAGATCTCGAGTCAATGCGGATGATGATG 4217
Qy 4202 GAACTCGATTTCTAAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 4261
Db 4218 GAGCTGGATATCAAAATACGCAAAATCTCAAGATGATTTCCACTGTGCTAGTAAATAC 4277
Qy 4262 GAAATATGAAAGATTTGGTCTCAATGATTTTGGCCGAGTGTGGAAACAAAGGAC 4321
Db 4278 GAGATCTGGGCAAGATTTGGGTTTGAAGACTTCTTGGGAGAGTTTGGAAACAAAGG 4337

Qy 4322 AGGAAAAAATTTTGAAGGATTAATCTGCTGGAATCAAGACATGTTCTGTGTATCAAGG 4381
Db 4338 AGAAAGACACCCCTCAGGATTAATCCGACAGGTATAAAACTTGCATCTGTATCAAGA 4397
Qy 4382 AAAAGCGGTGATGATGATCTTTCAATCGGCAATCTGTTTAAATAGAGCTGTGCTGG 4441
Db 4398 AAGAGCGGACGTCACGACGTTCAITGGAAACACATGATGATGATGATGATGATG 4457
Qy 4442 TCAATGTTACCGATGGAAGGTCATAAAAGGTCCTTTTGTGGAGACGATTCGTTTGTG 4501
Db 4458 TCGATGCTTCGATGGAAGAAATTAATCAAGAGGCTTTTTCGGTGACGATGATGCTGT 4517
Qy 4502 TATTTTCCAAAGGTTTGGATTTCCCTGACATTCAGTCAATGCTGCTAAATCTCAATG 4561
Db 4518 TACTTTCCAAAGGTTTGTGATTTCCGATGTCACACCTCCGCAATCTTATGCGAAT 4577
Qy 4562 TTTGAGGCCAAATCTGATAGAAAGGATACGTTTCTTTTGTGTAGATACATACAC 4621
Db 4578 TTTGAGCCAAATCTGTTTAAAAAACAGTATGATGATGATGATGATGATGATGATG 4637
Qy 4622 CATGATAGGAGCAATAGTATGATGATGATGATGATGATGATGATGATGATGATGATG 4681
Db 4638 CACGACAGAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 4697
Qy 4682 AAAATATCAAGGATTAATGATCACTTAGAAGATTAAGGTTCTTTTGTGCGATGTTG 4741
Db 4698 AAACATCAAGGATTTGGAAACCTTGGAGGATTCAGAAAGTCTCTTTGTGATGTTG 4757
Qy 4742 TGTTCGCTCGGAACTGTTAGGCTTCCGAGCTGACGCGATGATCAAGAGGTT 4801
Db 4758 GTTTTCGTT --- GAAACAAATTTGCGTATTTACACAGTTTGGACGACGCTGTA 4814
Qy 4802 CATAAAAACCGGATTTGATGTTTCTGTTTAAATTTGTTTAAACAAATTTTGTGAT 4861
Db 4815 CATGAAGCCGCTCCAGGTTCTGTTTGTATTAAGTCTGTTGAGATTTTGTCTGAT 4874
Qy 4862 AAATTTTATTTAAGACTTTTGTAAATGCTGTTAGTCTCAGAGATCTGTCAAAAT 4921
Db 4875 AAGTTCTTTTAGAAGTTTCTTTATAGATGCTAGTTTGTAAAGGAAAGTGAATAT 4934
Qy 4922 TAGCGGTTCAATCTTTTCGAAACAGGATGATGATCTCCGCGATTCATGACTAAG 4981
Db 4935 CAATGAGTTTATGACCTGACAAATAATGGAAGATCTTACCGTCAATGTTTACCC 4994
Qy 4982 CAAGAGTTTGAATATCGACTGTGGCAAGATTAATGCTGTTAAGAAATGATGCTTTC 5041
Db 4995 AAGAGTTTATGTTTCCAAAGTTGATAAATAATGTTTCTCATGAGATGATGATGCT 5054
Qy 5042 TGATGATGATTTAATAAGGTTTAAAGTTAGTTAAGAAAGGTTATGTTGCTTAGCTGA 5101
Db 5055 AGAGTGAACTCTTTTAAAGGATTTAAGCTTTATGATAGTGGATACGTTCTTTTAGC 5114
Qy 5102 TTTGTTAGTGTCTGGGAGTGAATCTCCGATAACTGCGGTGTTGTTGTTGTTGTTG 5161
Db 5115 TTTGTTGCTCAGCGGAGTGAATCTTCCCTGACAAATTTGAGAGAGGTTGAGGCTGTG 5174
Qy 5162 TATTGATAGTGAAGTGAAGAGGATTAAGAAAGCAACGCTGGGTGCTGATCAACGCCC 5221
Db 5175 TCTGTTGCAAAAGGATGGAAGAGAGCCGACGAGGCCACTCTCGGATCTTTACTACAG 5234
Qy 5222 TGCTTGGAAAGAAATTTTCTTTTAAAGTAAATCCCTAAATTTTCAATTAACATCCG 5281
Db 5235 AGTTCAAAGAAAGATTTTCAAGTTCAGGTTCCAGGTTCCCAATTTATGCTATTAAC 5294
Qy 5282 TGCTGAGAGCAACCGCTGCGAAGTGTAGTGAATCAAGAGGTTGCTTATGGAAGAG 5341
Db 5295 CCGGATGAAGAAACGCTCTGCGAAGTTTGTATTAATTAAGAAATGGAAGTGTGAG 5354
Qy 5342 ATACTGCTCTTTTATCTTTGGAGTTCGTTTCAATTTTGTGTAGTACATAAAATTA 5401
Db 5355 TTTCTGCTCGCTTTCTCTGGAGTTTGTGCTGCTGTTGTTTATAGAAATTAATATAA 5414

```
Qy 5402 AAAAGCTTTGAGGAAAGTATTTTGTAGTGTGACAGCGGCTCGCAATTGAACTCACTGA 5461
Db 5415 ATTAGGTTTGAGAGAGAGATTTACAAACGTGAGAGACGAGGCGCCATGGAACATTACAGA 5474
Qy 5462 AAAGGTTTGTGAGAGGTTGTTGATGAGTGAAGTACCAATGGCTGTGAAATCGGAAAGGTTTC- 5520
Db 5475 AGAGTGTGATGATGTTTCATGGAAGATGTCCTTATGTCGATCAGGCTTGCAAGTTTCG 5534
Qy 5521 -----CGGAAACAAAGAAAGGTTAGGTGATTAATGTTTAAAT----- 5561
Db 5535 ATCTCGAACCGGAAAAAGAGTGATGTCGCAAGGGAAGAAATAGTAGTAATGATCGTGC 5594
Qy 5562 -----RAGAAATTAATACATGTTGTAAGAGGTTTAAATTTAGGAAATGAGGA 5614
Db 5595 AGTCCGAAACAGAACTATAGAAATGTTAAGGATTTTGGAGGAATGAGTTTAAAAAGAA 5654
Qy 5615 TAATGTAAGTATGATCAGAGT-----CTATCGGTCATCGAGTACGTTTAAATCAAT 5665
Db 5655 TAAATTAATCGATGATGATTCGGAGGCTACTGTGCGCAATCGGATTCGTTTAA--AAT 5711
Qy 5666 ATGCTTTATACAACTCTCCGAGGCAATTTGTTTACTTATCTTCGCTTACGCGAT 5725
Db 5712 ATGCTTACATGATCACTACTCCATCTCAGTTCGTGTTCTGTGTCATCAGCGTGGCGGAC 5771
Qy 5726 CCTGTCAGCTGATCAATCTGTGTACAAATGTCATGTTGGTAAACGATTTCAACGCAACAA 5785
Db 5772 CCAATAGAGTTAAATTAATTTATGTAATGTAATCCCTTAGGAATTCAGTTTCAAAACAA 5831
Qy 5786 GCTAGGCAACAGTCCCAACAGCAATTTGCGGATGCTGGAACCTGTGCTAGTATGACA 5845
Db 5832 GCTCGAATCTGCTTCAAGACAATTCAGTGAGGTGTGGAACCTTCACCAAGTAATCT 5891
Qy 5846 GTGAGATTTCTGATCGGATTTCTATGTGTATGATATATAATTCGAGCTTTGATCCGTTG 5905
Db 5892 GTTAGGTTCCCTGACAGTGACTTTAAAGGTGTACAGGTACAAATCGGTATTAGACCGGCTA 5951
Qy 5906 ATCAGCGGTTTAAATAGCTTTGATGATCTAGAAATAGAAATAGAGGTTGATTAACAA 5965
Db 5952 GTCAAGCACTGTTAGGTGCAATTCGACACTAGAAATAGAAATAGAGGTTTGAAGATCAG 6011
Qy 5966 CCGCACCGAATCTACTGAAATCGTTAAACCGACTCAGAGGTTAGAGTGTCTACTGTGA 6025
Db 6012 GCGAACCCGACGCTGCGGAAACGTTAGATGCTACTCGTAGATGATAGCAGCGAACGGTG 6071
Qy 6026 GCTATAAGGGCTTCAATCAATTAATTTGCTTAATGAACCTGTTGCGGAACTGGCATGTT 6085
Db 6072 GCCATAAGGAGCGGATAAATAATTAATAGTAGAAATTTGATCAGAGGAACCGGATCTTAT 6131
Qy 6086 AATCAGCAGGCTTTGAGACTGTAGTGACTTGTCTGGACCAACATCCCGGCTACTTAG 6145
Db 6132 AATCGGAGCTCTTTGAGAGCTCTTCTGTTTGGTTTGGACCTCTGGTCTCTGCAACT--- 6188
Qy 6146 CTATTGTTGTGAGATTTCTTAAATAAAGTCGTGAAGACTTTAAATTCAGGGTGGCTGA 6205
Db 6189 -TGAGGTAGTCAAGATGATTAATAATTAACGATTTGTGCTGATATCA- -CGTGGTGGC 6246
Qy 6206 TACCAAAATCAGCAGTGTGTTGTTGCTCCACTTAATAATATAACGATTTGTCATATCTGGATCC 6265
Db 6247 TACGATAACGATAGTGTGTTTCCCTCCACTTAATAATCGAAGGTTGT- -GTCTTGGATCGC 6305
Qy 6266 AACAGTTAAACATGTGATGTGATGATGATGATGATGATGATGATGATGATGATGATGATG 6325
Db 6306 GCGGGTCAAAATGATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 6365
Qy 6326 AATCCTCCCTAAACCGCGGTAGGGGCCCA 6355
Db 6366 AATCCCCCGGTTACCCCCGGTAGGGGCCCA 6395
```

RESULT 7

US-09-259-741-2

; Sequence 2, Application US/09259741

; Patent No. 6033895

```
GENERAL INFORMATION:
APPLICANT: GARGER, STEPHEN
APPLICANT: HOLTZ, R. BARRY
APPLICANT: MCCULLOCH, MICHAEL
APPLICANT: TURPEN, THOMAS
TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
TITLE OF INVENTION: PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES FROM PLANT
TITLE OF INVENTION: SOURCES
NUMBER OF SEQUENCES: 5
CORRESPONDENCE ADDRESS:
STREET: 1299 Pennsylvania Avenue N.W.
CITY: Washington
STATE: DC
COUNTRY: USA
ZIP: 20004
COMPUTER READABLE FORM:
MEDIUM TYPE: Diskette
COMPUTER: IBM Compatible
OPERATING SYSTEM: DOS
SOFTWARE: FastSeq for Windows Version 2.0
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/259,741
FILING DATE: February 25, 1999
CLASSIFICATION:
PRIOR APPLICATION DATA:
APPLICATION NUMBER: 09/037,751
FILING DATE: March 10, 1998
ATTORNEY/AGENT INFORMATION:
NAME: Halluin, Albert P.
REGISTRATION NUMBER: 25,277
REFERENCE/DOCKET NUMBER: 00801.0140.US01
TELECOMMUNICATION INFORMATION:
TELEPHONE: 650-463-8100
TELEFAX: 650-463-8400
TELEX:
INFORMATION FOR SEQ ID NO: 2:
SEQUENCE CHARACTERISTICS:
LENGTH: 6439 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: unknown
MOLECULE TYPE: Genomic RNA
US-09-259-741-2
```

Query Match 38.6%; Score 2455.4; DB 3; Length 6439;

Best Local Similarity 44.8%; Pred. No. 0;

Matches 2772; Conservative 1147; Mismatches 2206; Indels 60; Gaps 8;

```
Qy 20 ACAACAACAATTAAACAAACAAACAAACATATTTACAAACAAACAAACAAACAAACAAACAAAC 79
Db 18 ACCAACCAACAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAACAAAC 77
Qy 80 ATACAATCTATATTAGCAACGCCCTTCTTGAAGCGTGAGTGTGTAACAAACACTCTCGTT 139
Db 78 ACACAGACAGCAACCAACAGCAACGCGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAGGAG 137
Qy 140 AATGACCTTGCAAGAGCGCATGTAGATACGCGCGTGGAGAAATTTAAACGCGCGAC 199
Db 138 AAUGAUCAGCAAGCGGCGUUCUUAACGACACGCGGUGAAGAGUUAUUAACGCGGAGGAGGAG 197
Qy 200 CGTAGACAAAGGTCAACTTTTCCAAACACTATTAGCGAGAGAGCAACCGTCTTAGTCTCC 259
Db 198 CGCAGGCGCAAGGUGAACUUUUAACAAAGAAUAAUAGCGAGGAGGAGGAGGAGGAGGAGGAGG 257
Qy 260 AACCGGTACCGGAGTTCAGATTACCTTTTATATCTCAAAATCTCAAAATCTCAAAATCTCA 319
Db 258 CGGCGGUAUCAGAAUUCGAAUUAACUUUAACGCGAAUUAACGCGAAUUAACGCGAAUUAACG 317
Qy 320 GCTGAGGTTTGAGAGCAATTAGAAATGGAATATCTGATGCTACAAAGTTCCCTATGATCG 379
Db 318 GCAGGUGAUGCGAUCUUAAGACUGGAAUUAUGAUGAUGAUGAUGAUGAUGAUGAUGAUGAUGA 377
```



```

; TOPOLOGY: unknown
; MOLECULE TYPE: Genomic RNA
; US-09-037-751-2

Query Match      38.6%; Score 2455.4; DB 3; Length 6439;
Best Local Similarity 44.8%; Pred. No. 0;
Matches 2772; Conservative 1147; Mismatches 2206; Indels 60; Gaps 8;

Qy 20 ACAACAACAATTAAACAAAAAACAACATATTTACAAACAACAACAACAACAATGGCACAC 79
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 18 ACCAACAACAACAACAACAACAACAACAACAUAUACAAUUAUUAUUAUUAUUAUUAUUAU 77

Qy 80 ATCAATCTATATTTAGCAACGCCCTTCTTGAAGCGTGAGTGGTAAAAACATCTCGTT 139
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 78 ACACAGACACUACCAACAUCAUCUUGUCGACACUGUCCGAGAAAACAACUUCUGGUC 137

Qy 140 AATGACCTTCGAAGAGCGCATGTACGATACCGCCGTGGAAGAAATTTAACGCCCGCGAC 199
   . ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 138 AUAUGAUCUAGCAAGCGUCGCUUUUACGACACACAGCGGUGAAGUUAUUAACGUCUGUAC 197

Qy 200 CGTAGACCAAGGTCCTTTTCCAAATCTATTAGCGAAGAGCAAAACGCTTTAGTCTCC 259
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 198 CGCAGGCCCAAGGUAACUUUUCAAAAGUAAUAAAGCGAGGAGCAGCGCUUAUUGCUACC 257

Qy 260 AACCGCTACCGGAGTTCAGATTACCTTTTATAATCTCAAAATGCCGTACACAGTTTG 319
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 258 CGGCGTGAUCCAGAAUUUCCAAAUUAUUUAUUUAACAACGCAAAAUUGCGUACGCUU 317

Qy 320 GCTGGAGGTTTGAGAGCATTTAGAAATTCGGAATATCTGATGTACAAGTTCCTATGATCG 379
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 318 GCAGUGGAUUGGAUCUUUUGAAACUGGAUAUUCUGAUGAUGCAAAUUCUUCGGAUCA 377

Qy 380 CCGACATATGATATAGGTGGGAACTTTGACGACATTTGTTTCAAAGCAGGGATTACGTG 439
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 378 UUGACUUAUGACAUGAGCGGGAUUUUGCAUCGCAUCUGUUAAGGAGCAGCAUAUGUA 437

Qy 440 CATTGCTGTATGCCAATCTGGACATACGAGATATATATGAGGACGAGAGCAAAAAGGAC 499
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 438 CACUUGUGAUGGCCAACCUUGGACGUGUUGGACAUCAUGCGGACGCAAGGCCGAGAAGAC 497

Qy 500 TCAATTGAGATGTTTGTGCCAGATTGTCTCGTTCTAAACAGGTAAATTCCTGAGTTTCAA 559
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 498 AGUUAUGAACUAUACUUUUCUGCUUAGGAGAGAGGGGGGAAAAACAGUCCCAACUCCAA 557

Qy 560 AGGGAGCTTTTAAACAGGTATGCAGAGCTCCCAAACGAACTCTGCTCTCTAAAACTTTT 619
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 558 AAGGAGCAUUUGACAGAUACGAGAAAUUUCUGAAGACCGUCUGUACAUUAUUAUUC 617

Qy 620 CAGATTGTGCAATATATCCGCCAGAGATAGTGGTGAAGATACGCTGTGCTCTGACAC 679
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 618 CAGACAAGCGACAUCAGCCGAUGCAACAUACGAGCAGAGUAGUUGCAUUGCGCUACAC 677

Qy 680 AGTTTGTATGATATCTCTGTGATGATGTTTGGAGCTCGGTTAATATCTAAGATATACAT 739
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 678 AGCAUAUAGACAUAACAGCCGAUGAUGUUGCGGGCGGCACUUGAGGAAAAAUGUCCAU 737

Qy 740 GTATGTTATGCAGCTTCCATTTTGGCAGAAGCATTTACTAGACACAGACGGAGTTTACG 799
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 738 ACGUGCUAUGCCGCUUUCACUUCUCCGAGAACUUGCUUUCUGAAGAUCAUACGUCAU 797

Qy 800 CTTAAATGAATATGGCGCAACTTTCAAAGAGAAAGGTGATGTTTCTTTTCTTTTGT 859
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 798 UUGGACGAAAUCAACGGUGUUUUUUCGCGAUGGAGACAAGUUGACCUUUUUCUUUGCA 857

Qy 860 GATGAAGTACTTTTAATTAATGATCAATAATAAATAATCTTGCATATTAGTTAAA 919
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 858 UCAGAGAGUACUCUAAAUUUAUGTCAUAGUUUAUUUCUAAUAUUCUUAAGUAGUGUGCAA 917

Qy 920 TCTTACTTTCCTGCTCTCTAGTAGAATAGTTTCTTTTAAGAAATTTTTTACTAGGTT 979
   ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 918 ACUUAUUCUCCGGCCUUAUAGAGAGUUUAUUAUGAAGAGGUUUUUUUAUGUACACAGUA 977

Qy 980 AATACTTTGGTTTTGTAAATTTTACAAAGTAGATACCTATATCTTGTAACAAGAGTTTGA 1039

```

[illegible]

4262 GAAATATGAAAAAGATTGGGTCTCAATGAGTTTGGCCGAAGTGTGGAACAAGGGCAC 4321
4278 GAGAUUCUGGGAAGUUGUUUUGAAGCAUUUUUUUUUUUUUUUUUUUUUUUUUUUUUU 4337
4322 AGGAAAAACAATTTGAGGATTTACATCTCTGGAATCAAGACATGCTGTGTATCAAAAGG 4381
4338 AGAAAGACCACTTCCAGAGUUUUAUCCGAGUUUUAUAAACUUGCAUCUGGUUUAAGA 4397
4382 AAAAGCGGTGATGACTTCTTCAATCGGCAATCTGTTAATAATAGCAGCTTGTGGGT 4441
4398 AAGAGCGGGACGUCACGACGUCUUAUUGGAACAACUGUAUCUUGCGCAUGUUGGCC 4457
4442 TCAATGTTACCGATGGAAGGTCATATAAGGTGCTTTTGTGGAGAGATTCGTTTGG 4501
4458 UCGAUGUUCUGAUGGAGAAAUAAUCAAAGGACUUUUUGCGGUGAGUAGUCUGUCU 4517
4502 TATTTCCTCAAGGGTTTGGATTTCCCTGACATTCAGTCATGCTGTAATCTCATGTGGAAT 4561
4518 UACUUCCTCAAGGGUUGAGUUUCCGAGUUGCAACACUCCGCGAUCUUAUGUGAAU 4577
4562 TTTGAGGCCAACTGTATAGAAAGGTPACGGTTACTTTTGTGTGAGATCATATACAC 4621
4578 UUGAAGCAAAACUGUUUAAAAACAGUAGGAUUAUUUUGCGGAAGAUUAGUAAUACAU 4637
4622 CATGATAGGAGCAATAGTATGATATGATCTTTGAAGTTGATCTCAAACTTGGGGCA 4681
4638 CACGACAGAGUAGUUGUUAUUAUAGUCCCUAAAGUUGAUCUGGAAACUUGUGUCU 4697
4682 AAACATATCAAGGATTAATGATCACTAGAACAGTTAAGGGTCTTTGTGCGATGTGCT 4741
4698 AAACACAAGUUGGGAACACUUGGAGGAGUUCAGAGUCUUCUUGUAGUUGUCU 4757
4742 TGTTCGCTCGGAACTGTGCTTGAAGCTTTCCGACGCTGAACGACGCTATCAAGAGGTT 4801
4758 GUUUCUGUAGAACAUUGUGGUA---UUAACACAGUUGGACGCGUUAUGGAGGU 4814
4802 CATAAACCGGATGATGTTGCTGTTTAAATGTTGTTAAACAAATTTTGTGTGAT 4861
4815 CAUAAAGCCGCCUCCAGGUCUGUUUUAUAAAGUCUGGAGUAGUUAUUGUCUGAU 4874
4862 AAATTTTATTAGAACTTTGTTTAAATGCTGTAGTCTCAGAGATCTGTCAAAAT 4921
4875 AAAGUUCUUUUAAGUUUUUUAAGUUGUCUUAUUGUUUAAGGAAAGUAGUAAU 4934
4922 TAGCGAGTTCAATGATCTTTGAAACAGGATGAGATCTTCGGCAITTCATGACTAAGT 4981
4935 CAUAGUUAUUCGACUUCUCAAAGGAGAGAGUUAUUAUCCGUGUUAUCCCGU 4994
4982 CAAGAGTTTAGAATATGATGCTGAGACAGATTAATGCTGTTAAGATGATGCTTTC 5041
4995 AAAGAGUUAUUGUUCUCAAAGUUAUAAUUAUUGUUAUUGGAGUAGUAGUUAUUG 5054
5042 TGATGTAGATTTACTTAAAGTGTGTTAGTTAGTTAAGAAAGGTTGTTGCTAGCTGA 5101
5055 AGAGUGAACCUUUUUAAGGAGUUAUUAUUGUUAUUGGAGUUAUUGUUAUUGGAG 5114
5102 TTTGTTAGTGTCTGGGAGTGAATCTCCCGATAACTCCGTTGTTGTTGTTGTTGTTG 5161
5115 UUGUGUCUACGCGGAGGAGGAACTUCUGACAAUUGCAGAGAGGAGGUGGUGUG 5174
5162 TATTGTAGATAGAGAAATAAGAGGATGAAGAGCAACGCTGGGTGTTATCAGGCCCC 5221
5175 UCUGUGGACAAAAGGAGUAGGAGGAGCGGAGGCGACUUCGCGAUUUAUUAUUAUUA 5234
5222 TGCTTGCAAAAGATTTTCTTTTAACTTAATCCCTAATTAATTAATTAATTAATTAAT 5281
5235 AGCUCAAAGAAAGAUUUGUUAUUGUUAUUGUUAUUGUUAUUGUUAUUGUUAUUG 5294
5282 TGCTGAGAGCAACCGGCTGAGTGTAGTGAATCAAGAGGTTGCTATGGAAGAAGG 5341
5295 CGCGAUGAAACCGUCGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 5354
5342 ATACTGCTCTTATCTTTTGGAGTTCGTTTCAATTTGTGTAGTATCAATAAAATAATGTAA 5401

5355 UUCUGUCGCUUUCUGGAGUUUGUCGUGUUAUUGUUUAUAGAAAUUAUUA 5414
5402 AAAAGGTTTGAAGGAAACGATTTTTCAGTGTGACAGACGCGCTCGCAATTTGAATCACTGA 5461
5415 AUUAGUUUGAGAGAGAAGAUUAACAACGUGAGAGCGGAGGCGCCAUUUAACA 5474
5462 AAAGGTTTGAAGGATTCGTTGATGAAGTACCAATGCTGTGAAATCTGCAAAAGTTTCC 5521
5475 AGAAGUCGUGUAGUUAUUGAAGUUGUCCUUAUGUCGAGUAGGCGUUGCAAAUUGC 5534
5522 GG-----AAAACAAAAGAAATGTTAGGTAAATATGTTTAATAAT----- 5561
5535 AUCUGAACCAGGAAAAAGAGUAGUCCGCAAGGGAUUAUUAUUAUUAUUAUUAUUA 5594
5562 -----AAGAAAAATAATTAACAGTGTGTAAGAGGGTTTTTAAAAATTGAGGAAATTTGAGA 5614
5595 AGUGCCGAACAAGAAUAUUAAGAAUUAAGAAUUAUUGGAGGAAUAGUUAUUAUUAUUA 5654
5615 TAATGTAAGTATGACAGAT-----CTATCGCTCATCGAGTACGTTTAAATCAAT 5665
5655 UAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 5711
5666 ATGCTTATATCAATCACTCTCCGAGCAATTTGTTTACTTATCTTCCGCTTACGAGAT 5725
5712 AUGUCUUAACAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 5771
5726 CCTGTGACGCTGATCAATCTGTGTAACAATGCAATTTGGTAAACAGTTTCAAAACGAAACA 5785
5772 CCAUAGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 5831
5786 GCTAGGACAAACAGTCCCAACAGCAATTTGCGGATGCTGGAACCTGTGCTGTATGACA 5845
5832 GCUCGAAACUGUCGUTCAAAGCAAAUUCAGUGAGGUGUGGAAACCUCAACCAAGUAAU 5891
5846 GTGAGATTTCTGCTGATCGGATTTCTGTGTATAGATATAATTCGACGCTTGATCCGTTG 5905
5892 GUUAGGUUCCUGACAGUGACUUUAAGUGUACAGGUACAAGUUGGUAUUAUUAUUAUUA 5951
5906 ATCAAGCGGTTTATTAATAGCTTTGATAGTAAATAAGAAATAAGAGGTTTGAATCA 5965
5952 GUCACAGCAGUUAUGGUGCAUUCGACACUAGAAUUAAGAAUUAUUAUUAUUAUUAUUA 6011
5966 CCGCAGCAGAAATCTACTGAAATCGTTAAACGAGTACAGAGGTAGACGATGCTACTGA 6025
6012 GCGAACCCAGCAGTCCGAAACGUAUGAUGUACUACUGUAGUAGUAGACGACGCAACG 6071
6026 GCTAAGGGCTTCAATCAATAATTTGCTTAATGAACTGTTTCTGTTGAACTGGCATGTT 6085
6072 GCCAAGAGGAGCGGUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 6131
6086 AATCAAGCAGGCTTTGAGACTGCTAGTGGACTTGTCTGAGCCACCACTCCCGCTACTTGA 6145
6132 AAUCGAGCUCUUCGAGAGCUCUUCUGUUGUUGUUGGAGCUCUGGUGGCAACCUAG 6191
6146 CTATT 6150
6192 CAUU 6196

RESULT 9

US-09-466-422-2
; Sequence 2, Application US/09466422
; Patent No. 6303779

GENERAL INFORMATION:

APPLICANT: GARGER, STEPHEN
HOLTZ, R. BARRY
MCCULLOCH, MICHAEL
TURPEN, THOMAS

TITLE OF INVENTION: A PROCESS FOR ISOLATING AND
PURIFYING VIRUSES SOLUBLE PROTEINS AND PEPTIDES
FROM PLANT SOURCES

NUMBER OF SEQUENCES: 5

QY	1100	AAAACTTGGCCATGTTCAACACTGAAGAGACAATCTTTTAGACACACCGCTTCGGTTAAAC	1159
DB	1098	AAGACUCUUGCAUUGUGCAACAGCGAGAGAUAUCCUUCUGAUGAUUCAUCAUGACUACA	1157
QY	1160	TTTTTGTTCCCTTAAGATGAAGAGCATGGTGATAGTACCCTGTGTTGAGGGTTCTATTACC	1219
DB	1158	UACUGGUUUCCCAAAUGAGGGAUUGGUCAUGUACCAUAUAUCGACAUAUUUUUUGGAG	1217
QY	1220	AGCAAAGAAGATGACAAAGGAGTGAAGTCAATGTTAAATCGTGACTTCGTTTATACAGTGCT	1279
DB	1218	ACUGAUGAGAGACCGCRAAGCUUUGUCCUCCAAGAUUUUGUGUUAUACAGUGCUU	1277
QY	1280	AATCATATCAGAACAFATCAAGCCAAAGCGTTAACTTACAGAACGTAATATCTTTTCGTG	1339
DB	1278	AACCAAUUCGAACAUAACACGAGCGAAAGTCUUUACAUAUCGCAAAUGUUUUUCCUUG	1337
QY	1340	GAGTCTATAAGATCCCGCGTGATAATCAATGGTGTACTGCTAGGTCTCAATGGGATGTA	1399
DB	1338	GAAUCGUAUCUGGAGGGUUAUCAUUNACGGUGUACAGCGAGGUCGGAUUGGAUGUG	1397
QY	1400	GATAAAGCAATCTTTCAACCCTTGTCAATGACTTTTCTTCTTGAGACTAAGCTGGCTGGC	1459
DB	1398	GACAAUCUUUUAACAUCUUGUCUAGCACGUUUUACUGCAUAUCAUAGCTUGCCGU	1457
QY	1460	CTTCAAGACCATATAGTAATGCGGAAGATTTTCGGTGTGTAAGACCACTTCTGAACTT	1519
DB	1458	CUAAGAUGACUUAUCUGAUUAGCAUUAUUGUCUUGUUCGUAACGGUGGCCAGCAU	1517
QY	1520	ATTTGGGATGAGGTGGGCAAAATTTTTTGGAAACGTTTTTCCCACTATCAAGAAGATTG	1579
DB	1518	GUGUGGACAGAUUUCUGCGUGUUGGAAACGCAUUUCCUCCUGGAGAAAGAGACUC	1577
QY	1580	GTGAGCAGGAAATCTGSGATGTAAAGTGAGAAATGCTGGAAGATCAAGATCCAGATCTG	1639
DB	1578	UUGAACAGGAAACUUUACAGUGGCAGGCGACGCAUTAGAGAUCAAGGUGUCUGAUUA	1637
QY	1640	TATGTCATCTGMAAAGACAGGTCGTAGCTGAATATACACCAAGTCTGAGGATTTACCGCAT	1699
DB	1638	UAUGUGACCUCCACGACAGAUUAGUCAGUACAAGGCCUCUGUGGACAUUGCUGCG	1697
QY	1700	CTAGATATCAAGAAGACTTAGAGAGCTGAGCAATGTACGACCGGTTATCAGAAATTA	1759
DB	1698	CUGACAUUAGGAAGAAGAUGGAAGNAACGGAAGUGAUGAUCAUUGCAUUUUCAGAGUUA	1757
QY	1760	TCTATCTTAAAGGTGCTGTAATTTTCGATATCCGGAAGTTCAAAGACATGTGCAAGGCT	1819
DB	1758	UCGGUUAAGGGAGUCUGACAAAUUCGAUGUUAUGUUUUUCCCCAGAUGGCCAAUCU	1817
QY	1820	TTAGATGTTAGTCTGATGTGGCAGCAGGAGTATCGTTGAGTGGCCGAGATAGNAGC	1879
DB	1818	UUGGAAGUUGACCCAAUAGCGCAGCGAAGGUUAUUGUCGCGGUCAUGACGAUAGAGAC	1877
QY	1880	GGTTTAACCTCTACTTTTGAATGCAACCGAGGAGAAATGTGGCTAAAGGCTCT-----T	1933
DB	1878	GGUUGACUCUCAUUAUGACGACCUACUGAGCGGAUUGUGCGUAGCUUUAACAGGAU	1937
QY	1934	AAAAGCACGGCGCTGAGGCCGTGGTATGTTCTTGAAACCGACATCCGAAGAGGTGAACGTA	1993
DB	1938	CAAGAGAAGGGCUUCAGAAGGUGCAUUGGUAGUUAUCCUACAGAGAAGTUGAAGAACCGUCC	1997
QY	1994	AATAAATTTCTATTGCTGAGAAAGGGAGATTGCTGTGTGTCAGAAAAGTCATGGTTTG	2053
DB	1998	AUGAAGGUUCGAGUGCCAGGAGAGAUUAUCAAUUGUGUGUCUUGUGAGAUCAUCCG	2057
QY	2054	ACGAATGCTTAACCTTAGAGCAACGAGGTTGGAGTCCCTCAACGATTTTCATAAGGCTTGC	2113
DB	2058	GAGUCGCUUAUUCUUAAGAACGAGAGAVAGAGUCUUUAGAGCAGAGUUAUUAUUGCCACG	2117
QY	2114	GTGGATAGTGATTTACAAAGCAAAATGCGTTGCTACACTGGCTCACTCAAAAGTT	2173
DB	2118	GCAGAUCCGUUAUUGACGAGAGUCGAGUCGUAUGUGUUAUACGCGGUCGUAUUAAGUU	2177
QY	2174	CAACAAATGAAGAACTATGTGGACAGTTTGGCAGCTTCGTTGTCGCGCACTGTATCAAT	2233

Db	2178	CAGCAAAUGAAAAACUUUAUCGUAAGCCUGGUAGCAUCAUCAUCUGUCGGUGCGUAU 2237
Qy	2234	CTATGCAAGTCACATAAAGGATGAAGTCGGGTATGATTCGTATTCAGGGAGAAAGTTGGT 2293
Db	2238	CUCGUCAGGAUUCUCAAGAUACAGCUGCUUAUUGACCUUGAAACCCGUCAAAAGUUUGGA 2297
Qy	2294	GTTTGGGATGTCACCTTTGAAAAGTGGCTCTCAAACTGCGCGCCCAAAAGGTCATTCATGG 2353
Db	2298	GUCUGAGUUGUUGCAUCUAGGAAGUGGUUAUCCAAACCAACGCCCAAGUCUGCAUGG 2357
Qy	2354	GGAGTCTCTCGATTACAAAGGGGAAAAATGTTTACTGCACCTTCTATCTTATGAAGGAGAT 2413
Db	2358	GGUUGUUUGAAACCCACGCGAGAGAUCAUGUGCGCUUUUGGAAUAUGAUGAGCGAG 2417
Qy	2414	AGAAUATGTGACTGAGAGCGACTGGAGAGGGTGGTGTATCATCTGTGATCATATGTTAT 2473
Db	2418	GGUGUGUGACAUUGCGAUGAUGGAAGUAGCUGUUGAGCUCUGAGUCUCUGUGUUUAU 2477
Qy	2474	TCTGATATTCGAAGCTCCAAAATCTGAGGAAACAAATGAGAGACGGTGAAACCCACGAA 2533
Db	2478	UCCGACAUGCGGAAACUCAGAACUCUUGCGCAGACUCUUGGAAACGAGAACCGCAUGUC 2537
Qy	2534	CCTACTGCAAGATGGTATCTGTGGATGGGGTCCTGGTGTGGGAAAGTACAAAGAGAT 2593
Db	2538	AGUAGCGCAAAGGUUGUUCUUGGCGGAGUUCGGGCGUGUGGAAAAACCAAGAAAUU 2597
Qy	2594	TTTGAAGATTTGATCTTGATGAGGATTTGATCTTGTTCTCGAAAAACAAGCTGCTGCT 2653
Db	2598	CUUCCAGGGUUAAUUUUGAUGAAGAACUUAUUUUUAGUACCUUGGAAAGCAAGCCGGAA 2657
Qy	2654	ATGATCAGAAGAGGGCTAATTCATCTGGACTGATTAAGAGCCCAATATGACAAATGTGAGA 2713
Db	2658	AUGAUCAGAGACGUGCGAAUUCUCAGGGAUUUUGUGGCCACGAAGGACAAACGUUAAA 2717
Qy	2714	ACGGTAGATTCCTTAAATGCATCCAAACCCGATCACA-----CAAGAGGCTT 2764
Db	2718	ACCGUAGAUUCUUCAUGAUGAAUUUUGGAAAGCACACGUCUGAUUUCAGAGUUUA 2777
Qy	2765	TTTATTGATGAAGGTTTGATGCTGCACACCGTTGTGTTAACTTCCTGGTGTCTTCTCT 2824
Db	2778	UUAUUGAAGAAAGGUUGAUGUUGCAUACUGGUUGUUUAUUUCUGUGGCGAUGUCA 2837
Qy	2825	GGTTCCGACATCGCATACATTTACGGAGATACACAGCAGATTCCTTTCATTACAGATT 2884
Db	2838	UUUGCGGAAUUGUAUUGUUUACGGAGACACACAGCAGAUUCUCAUCAUAUAGAGUU 2897
Qy	2895	CAGAAATTTCCGTTATCCCAAACATTTTGAAGCTGCAAGTGGATGAAGTTGAGATGAGG 2944
Db	2898	UCAGAUUCCGUAACCCCGCCCUUUUGCCAAUUAUGGAUGUAGCGAGGUGGAGACACGC 2957
Qy	2945	AGGACCACTGAGATGCCAGGTGATGTGAAATTTTTTCTTACATCGAAGTACGAAGGA 3004
Db	2958	AGAACUACUCUCCGUUGCCAGCGAUGUCACACAUUUCGAAACGAGGAUUGAGGGGC 3017
Qy	3005	GGGGTGACAAACAATTCACTGACACGATCGGTCTCATCTGAGATGATAGCGGTAAG 3064
Db	3018	UUUGUCAUGACACUUUCUUGGUUAAAAGUCUGUUUCGACGAGAGUUGCCGCGAGCC 3077
Qy	3065	GGAGTACTAAACAGTGTGTTCCAAACCACTAAAGGGAAAAATGTGAACTTTTCACTCAGGCT 3124
Db	3078	GCCGUGAUCAUCCGAUCUCAAAACCCUUGCAUGGCAAGAUCCUGACUUUUACCCAUCG 3137
Qy	3125	GATAAATTTGATTAGAGAGAGGGCTATAAGATGTGACACCGTTTCATGATGATCCAA 3184
Db	3138	GAUAAAGAGCUCUCUUCAAGAGGGUUAUCAGAUUUCACUGUGCAUGAAGUGCGAA 3197
Qy	3185	GGAGAAACCTTTGAAGATGTGTGCTGTGATTCAGATTGACGGCAACTCCACTGACTCTGATT 3244
Db	3198	GGCGAGACADACUCUGAUGUUUCAUAGUAGGUUAACCCCUACACACAGUCUCCAUCAUU 3257
Qy	3245	TCCAAGTCTTCCCGCATGTTCTAGTCGCTGTGACTAGACACAAAGAGCTTCAATAT 3304

Qy	620	CAGGATTTGTCGAATACATATCCGCCAGAGAAATAGTGGTAGAAGATACGCTGTGCTCTGCAC	679
Db	618	CAGCAUUGGGCAUCAACGCGCAUGCAACAACAGSCAGAGUGUAUGCAUUTUGGCGUACAC	677
Qy	680	AGTTTGTATGATATTTCCCTGTGCATGAGTTTGGAGCTCGTTAAATATCTAAGAATATACAT	739
Db	678	AGCAUUAUGACAUACACAGCGAUGAGUUCGGGGCGGCACUUCUGAGGAAAAUUGUCCAU	737
Qy	740	GTAATGTTATGCAGCTTCCATTTTGGCAGAAAGCATTTATCTAGACCCAGACGGAGGTTACG	799
Db	738	ACGUGCUAUGCGCUUUCUUCUUGAAGAACCCUGCUUCUGAAGAAUUAUACUAGCUCAU	797
Qy	800	CTTAATGAATAAGCGCGCAACTTTTCAAAGAGAGAGTGATGATGTTTCTTTTCTTTGCT	859
Db	798	UUGACGANAUAACAACCGGUGUUTUUCGCGGAGUGGAGACAAGUUGACCUUUUCUUUGCA	857
Qy	860	GATGAAGTAGTACTTTAAATTTATAGTCATAAATACAAAAATATCTTGCATTTATGATTAA	919
Db	858	UCAGAGAGUAUCUUAUAUUAUGUCAUGUAUUCUUAUAUUCUUAAGUAUGUGGCAAA	917
Qy	920	TCCTTACTTTCCTGCTCTCTAGTAGAATAGTTTACTTTTAAGGAAATTTTACTTACGAGG	979
Db	918	ACUUAUCUUCGCGCCUUAUAUAGAGAGUUAUUAUGAAGAGUUAUUAUGUACACAGAGU	977
Qy	980	AATACTTGGTTTTGTTAAATTTAACAAGTAGATACCTATATCTGTGTACAAGAGTGTAGA	1039
Db	978	AAUACCGUGUUTUGUAAAGUUAUUCAGAAUAGAUACUUCUUCUUAUUAACAAAGGUG	1097
Qy	1040	CAAGTAGGTGTGATAGTGTACGTTCTATGAGGCGATGAAAGCGCTTGTCTTACAAG	1099
Db	1038	CAUAAAGUGUAUGAUGAGGAGAGUUAUAUCGCAUUGAAGACGCAUGGCAUUAACAA	1097
Qy	1100	AAAACTTGGCCATGTTTCAACACTGAAAGAGCAATCTTTAGAGACACGGCTTCGGTTAAC	1159
Db	1098	AAGACUUCGCAUGUGCAACACGCGAGAAUCCUCCUGAGGAUUAUCAUCAGUCAU	1157
Qy	1160	TTTTTGGTTCCCTAAGATGAAGGACATCGTGATAGTACCGCTGTTTGAAGGTTCTATTACC	1219
Db	1158	UAUCUGUUTUCCAAAUAUGAGGUAUUGGUAUCGUAACCAUUAUUCGCAUUCUUGGAG	1217
Qy	1220	AGCAAAAGATGACAAGGAGTGAGGTCATTTGTTAAATCGTGACNTTCGTTTACAGAGCTT	1279
Db	1218	ACURGAUAGAGGACGCGCAAGGAAGUCUUAUGUGUCCAAAGGAUUAUCGAGUCCU	1277
Qy	1280	AATCATATCAGAAATATCAAGCCAAAGCGTTAACTTTACAGAACGCTATTATCTTTCGTG	1339
Db	1278	AACCAUAUCGACAUACACGCGGAAGCUCUUAUCAUCGCAAAUUAUUGUCCUUGUC	1337
Qy	1340	GAGTCTATAAGATCCCGCGTGATAATCAATGGTGTTACTGTAGGTCGAATGGGATGTA	1399
Db	1338	GAAUCGAUUCGACUCGAGGUAUAUCAUUAACGUGUGACAGCGAGGUCGGAUUGGAGUG	1397
Qy	1400	GATAAAGCAATCTTCAACCTTGTCAATGACNTTCTTCTTGCAGACTAGCTGGCTGCG	1459
Db	1398	GACAAUCUUGUUAACAUCUUGUCUAGACGUAUUCGUAUUCGUAUUCGUAUUCGUAU	1457
Qy	1460	CTTCAAGACGATATAGTAATGGGAAAAGTTTCGGTGCTTGGATAAGACCACTTCTGAAC	1519
Db	1458	CUAAAGAGUACUUAUGAUAUUGCAUUAUGUUCUGUUGCAAAACGUGUCCAGCAU	1517
Qy	1520	ATTTGGGATGAGGTGGGCAAAATTTTGGAAAAGTTTTTCCCCACTATCAAAAGAGAGATG	1579
Db	1518	GUGUGGUAUGAUUUCGUGCGGCUUUGGAAACGCAUUCUCCUGGUAAGAGAGGCGUC	1577
Qy	1580	GTGAGCAGGAAAATCTTGGATGTGAAGTAGAATGCTCTGAAAGATCAAGATCCCGATCTG	1639
Db	1578	UUGAACAGGAAAACUUAUCAGAGUGGCAAGGCGGACGCAUUAAGAGAUACAGGUGCUGA	1637
Qy	1640	TATGTACATGGAAGACGTTCTGTAGCTGAATAACCAAGTCTGAGGAGTTTACCGCAT	1699
Db	1638	UAUGUACCUUCCACGACAGUAUUGACUGAUAAGGCGUCCUUGUGGACUUGCCUUGC	1697
Qy	1700	CTAGATATCAAGAGGACTTAGAAGAGAGCTGACGAAATGTATGACGACGCTTATCAGA	1759

Db	1698	CUUGACAUUAGGAAGAAGUUGGAAGAAACCGAAGUAGUATACAAGUACAUUUCAGAGUUA	1757
Qy	1760	TCATATCCTTAAGGGTGTGTATAATTTTCGATATCGCGAAAGTTCAAAGACATGTGCAAGGCT	1819
Db	1758	UCGGUGUUAAGGAGUCUGACAAAUUCGAUUGUUGAUUUUUUCCAGAUUGGCCAAUCU	1817
Qy	1820	TTAGATGTTAGTCCTGATGTGGCAGCAGAGTAATCGTTGTCAGTGGCCGAGAAATAGAGC	1879
Db	1818	UUGAAUGUACCAACCAUAGACCGCAGCGAAGUUAUAGUCGCGUCAUGACCAUAGAGAGC	1877
Qy	1880	GGTTTAACCTCTACTTTTGTATAGCCAAACCGAGGAGNATGTGGCTAAGGCTCT- - - - -T	1933
Db	1878	GGUCUGACUCUCAUUUAGAACGACCUACUGAGCGGAUUGUUGCGUAGCUUUAACAGGAU	1937
Qy	1934	AAAAGCACCGCGTCTGAGGCGGTGTATGTTCTTGAACCGACATCCGAAGAGGTGAACGTA	1993
Db	1938	CAAGAGAGGCUUCAGAGGUGCUUUGUAGUUAUCUACAGAGAAGUUGAAGAACCGUCC	1997
Qy	1994	AATAAATTTTCTATGTCTGAGAAAGGAGGATTCGCTGTGTGTGTCAGAAAGTCAATGTTTG	2053
Db	1998	AUGAAGGUGUUGAUGGCGCAGAGGAGAGUUAACAUAUAGCUGGUGUCUGGAGAUCAUCCG	2057
Qy	2054	ACGAACTCTTAACCTTAGACACCGAGGAGTTGGAGTCCCTCAACGATTTCCATAGSCTTGC	2113
Db	2058	GAGTGCUGCCUUAUUCUUAAGAAACGAGGAGUAGAGUCUUUAGAGCAGUUAUCAUUGGCAACG	2117
Qy	2114	GTGGATAGTGTGATTACAAGACCAATGGCATCGGTTGTCTACACTCGCTCACATCAAAGTT	2173
Db	2118	GCAGAUUCGUUAUUCUUGAAGCAGUAGAGCUCUAUUGUUAACAGGUCGUAUUAAGUU	2177
Qy	2174	CAACAAATGAAGAACTATGTGACAGTTTGGCAGCTTCGTTGTCCGCCACTGTATCAAAAT	2233
Db	2178	CAGCAAAUGAAAAACUUUAUCGAUAGCCUGGUAAGCAUCACUAUCUGUGCGGUGUCGAU	2237
Qy	2234	CTATGCAAGTCACTAAAGGATGAAGTCCGGTATGATTTCTGATTTCCAGGGGAAAGTTGGT	2293
Db	2238	CUCGCAAGAUAUCUCAAGAAUAACGUCUGCUUUAUGACCUUGAAACCCGUCAAAAUUTUGGA	2297
Qy	2294	GTTTGGGATGTCACTTTGAAAAAGTGGCTCCTCAAACCTCGCGGCCAAAGGTCAATTCATGG	2353
Db	2298	GUCUUGAUGUGUACUUCUAGGAAGUGGUUAUCAAAACCAACGGCCAGAGUCUAGCAUGG	2357
Qy	2354	GGAGTTGTCCTGGATTACAAGGGGAAAAATGTTTACTGCACTTCTATCTTATGAAGAGAT	2413
Db	2358	GGUUGUUGAAACCCACGCGAGAGAAUAUCAUGUGGCGCUUUGGAAUAUUGAUGAGCAG	2417
Qy	2414	AGAAATGTGATCAGAGCGACTCGAGAGGGTGGTGTATCATCTGATACAAATGGTATAT	2473
Db	2418	GGUGUGUGGACAUCGCAUGAUUUGGAGAAGUAGUCUGACGUCUGAGUCUGUUGUUUAU	2477
Qy	2474	TCTGTATTCCAAGCTCCAAAAATCTGAGGAAAAACAATGAGAGACGGTGAACCCACGAA	2533
Db	2478	UCCGACUUGGCGAAACUCAGNACUUCGCGCAGACUGUCUUGAAACGGAGAACCGCAUGUC	2537
Qy	2534	CCTACTGCAAGATGTGTACTTGTGGATGGGGTGCTGGTTGTGGAAAGTAACAAGAGAT	2593
Db	2538	AGUAGCGCAAGGUGUUCUUGUGGACGGAGUUCUGCGGUCUGGGGAAAAACCAAGAAAAU	2597
Qy	2594	TTTGAAGAATTTGATCTTGTATGAGATTTGATCTTGGTTCTCGGAAAAACAAGCTGCTGCT	2653
Db	2598	CUUCCAGGGGUUAUUUUGAUGAAGAUCAUAAUUUUAUAGUACUCUGGGAGCAAGCGCGGAA	2657
Qy	2654	ATGATCAGAGAAGGGCTAATTCATCTGGACTGATGAAGGCCCAATGGCAATGTGTAGA	2713
Db	2658	AUGAUCAGAGAAGUGGGAUUUCUCAGGGAUUAUUUGGCGACAGAGGCAACACGUUAAA	2717
Qy	2714	ACGGTAGATTCATCTTAATGATCCAAAACCGCGATCACA- - - - -CAAGAGGCTT	2764
Db	2718	ACCGUUAUUCUUCAUGAAGAUUUUUGGAAAAAGCACCGUCUGAUUCAAGAGGUUA	2777
Qy	2765	TTTATTGATCAAGGGTTGATGCTGCACACCGGTTGTGTAACTTCTCGTGCTTATCTCT	2824

Db 2778 UUAUUGAUGAGGUGAUGAUGCAUACUGGUGUUGUAUUUUCUUGGCGGAUGUCA 2837
Qy 2825 GGTTCGCAATGCGCATATTAATTCAGGAGATACAGAGAGATTCCTTTCAITTAACAGAGTT 2884
Db 2838 UUGUGGAAAUUGCAUUGUUAUUCGAGACACACAGCAGAGAUUCCAUACAUAUGAGUU 2897
Qy 2895 CAGAATTTCCCGTATCCCAACATTTTCAGAGCTGCAAGTGGATGAATGAGATAGG 2944
Db 2898 UCAGAUUCCCGUAUCCCGCCCAUUCGCAUUAUUGCAUUAUUGAAGUUGAGAGGAGACACGC 2957
Qy 2945 AGGACACACTGAGATGCGCAGGTGATGTGAATTTTTCTCAATCGAAGTACGAAGGA 3004
Db 2958 AGAAUAUCUUCGUGUUGCCAGCCGGAUGUCAACAUAUUCUGAAACAGGAGAUUAGAGGC 3017
Qy 3005 GCGGTGACAAACACTTCAACTGTACACAGATCGGTCTCATCTGAGATGATAGGCGGTAG 3064
Db 3018 UUGUGAUGAGACAUUCUUGCGUUAUUAAGUCUGUUGCGAGGAUGUGCGGAGGCC 3077
Qy 3065 GGAGTACTAAACAGTGTTCCTAAACCCACTTAAAGGGAATTTGTAATTTCACTCAGGCT 3124
Db 3078 GCGGUAUACAUCCGAGUCCUAAACCCUUGAUGGCAUGAGUCCUGACUUAUCCCAUUG 3137
Qy 3125 GATAAATTTGAGTTAGAGAGAGGCTATAAGAAATGTGAACACACCGTTTCAATGAGATCCAA 3184
Db 3138 GAUAAAGAGCUCUGCUUAAGAGGUAUUCAGAUUGUCAACUGUGCAUGAAGUGCAA 3197
Qy 3185 GGAGAAACCTTTGAGATGTTCGCTGCTGAGATTCAGCGCAACTCCACTGACTCTGATT 3244
Db 3198 GCGGAGACAUACUCUGAUGUUCACUAGUUAUAGUUAUCCUUAACAGCAGUCCUUAUUC 3257
Qy 3245 TCCAACTTCTCCCGCATGTTCTAGTCTGCTGACTAGACACACAAAGAGCTTCAAAATAT 3304
Db 3258 GCAGGAGACAGCCCAUUGUUGUGCAUUGUCAAGGACACACUUGUCCUCAAGUAC 3317
Qy 3305 TACACCGTAGTGTAGATCTTTAGTACAGATAATTAAGTATGTTGTTCTTTTAAGTCC 3364
Db 3318 UACACUGUUAUGAUGAUCUUAUUGUAUUAUAGAUAGAUAGAGAAACUUAAGCUG 3377
Qy 3365 TTCTTTTGAATGTATGTAGAGAGCAGGTAGTAGATAGCAATTAACAGATGATGCA 3424
Db 3378 UACUUGUUAUGAUGUUAAGGUGGAGUAGCAGGAAACAAUAGCAUUAUAGAUAGCUG 3437
Qy 3425 GTGTTCAAGGTCAATCTCTTTGTGGCAACACCTTAATCAGGAGACTTCCAGATCTA 3484
Db 3438 GUGUCAAAGUUCCAUUCUUGUGCAGCCAAAGACUGGUGUAUUAUUCUGAUUG 3497
Qy 3485 CAGTTCTATTACATGTATGCTTCCCTGCTGTAAGTACTATCTTAAACAGTATGATGCT 3544
Db 3498 CAGUUAUUAUGAUGUGUCUCCAGGCAACAGCAUUAUAGAUUAUUAUUGAUGCU 3557
Qy 3545 GTTACCATGAGGTTACGTGATAATAGTCTTAATGTGAAGGATGTGTTCTTGAATTTTCC 3604
Db 3558 GUUACCAUGAGGUGUACUGACAUUUAUUAUUGCAUUAUUGCAUUAUUGGAUUGCU 3617
Qy 3605 AAAAGTATTCGATGCCAAAGGAGGTGAACACCAATGTCTAGAGCCAGTTTTCGCTACGGG 3664
Db 3618 AAGUCUGUGUGCGCCUAGGAUCAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3677
Qy 3665 GCGGAACCGCAAGGCTGACGACTACTCGAATCTCGTTCGATGATTAAGAAAT 3724
Db 3678 GCAGAAUUGCCACCCAGACUGGACUUAUUGCAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3737
Qy 3725 TTCAACGCACACACTGACGGGACGANTTGAATTCAGAGACACCGCATCTGTGTAGTA 3784
Db 3738 UUAACGCACCCGAGUUGUGUGCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3797
Qy 3785 GATAAGTTTTTGTAGTATTTTATTAAGAAAGAAATACACAAATAATTTCTGGA 3844
Db 3798 GAUAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3857
Qy 3845 GTGATGACGAAGGATTCATCATCATGATGTTTGAAGAAACAGGAAGAGTACTATTGGAC 3904
Db 3858 UUGUUAUGAUGAGUCUCUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3917

Qy 3905 GACTTGGCTAACTACAAATTTTACAGATCTGCGGCCCATCGATCAGTACAAAGCACATGATC 3964
Db 3918 CAGCUCGAGAUUUGAUUUGUAGAUUUGCCAGCAGUUGUAGUACAGACAGAUU 3977
Qy 3965 AAGGCTCAACCAAAACAGAAATTTGGACCTTTCAATTCAGAAATGAATACCTCTCTGCA 4024
Db 3978 AAAGCAAAACCCAAAGCAAAAUUGGACAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4037
Qy 4025 ACAATGTCTACCATTCGAAGCAGATCAACGCTATTTTGGCCGGTTTCT---CAGAGCTT 4081
Db 4038 ACAGUUGUGURCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4097
Qy 4082 ACAAGGTTGCTGCTCGAGGCAATTTGATTTCTAAGAGTTTCTTTCTTTTCTTACTAGAAACT 4141
Db 4098 ACUAGGCAUUAUCUGGACAGUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4157
Qy 4142 CCAGAACAGATTCAGAAATTTTCTCGGATCTGACCTCGACGTTCTCTATGGATGTGTA 4201
Db 4158 CCAGCGAGAUUGAGGAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4217
Qy 4202 GAACTGATATTTCTAAGTATGATGATCAAGAACGAGTTTCAATGTGTGTAGATAT 4261
Db 4218 GAGCUGAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4277
Qy 4262 GAAATATGGAATAAGATTTGGTCTCAATGAGTTTGGCCGAAGTGTGGAACAGAGGCAC 4321
Db 4278 GAGAUCUGGCGAAGAUUUGGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4337
Qy 4322 AGGAAACACTTTTGAAGGATTAATTTGCTGGAATCAAGACATGCTGTGTGTATCAAGG 4381
Db 4338 AGAAGAGCACCCCUAAGAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4397
Qy 4382 AAAACGCGTGTGTGACTTCTTTCATCGGCAATCTGTTATAAGTATTAAGTATGAGTGTGGT 4441
Db 4398 AAGAGCGGGACGUCACGACUUAUUGAAACACUGUGAUAUUAUUAUUAUUAUUAUUAUUAUUA 4457
Qy 4442 TCAATGTTACCGATGGAAGGTCATAAAGGTCCTTTTGTGGAGAGATTCGCTTTTG 4501
Db 4458 UCGAUGUUCGAGUUGGAGAAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4517
Qy 4502 TATTTTCCAAAGGTTTGGATTTCCCTGACATTCAGTCAATGTGTATCTCATGTGGAAT 4561
Db 4518 UACUUAUUAAGGUGUUGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4577
Qy 4562 TTTGAGGCCAACTGTATAGAAAGAGGTACCGTTACTTTTGTGTGATATCATATACAC 4621
Db 4578 UUAAGCAAAACUGUUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4637
Qy 4622 CATGATAGGAGCAATAGTGTATTCATCTTTTGAAGTTGATCTCCAAACTTTGGGGCA 4681
Db 4638 CACGACAGGAGUAGCAUUGUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4697
Qy 4682 AAAATATCAAGGATTTATGATCATTAGAAGATTTAAGGGTGTCTTTGTGGATGTTGCT 4741
Db 4698 AAACACUACAGAUUGGGAACACUUGGAGGAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4757
Qy 4742 TGTTCGCTCGGAAACTGTGTCTTACGCTTTCCGAGCTGAACGAGCTTATCAAGGAGTT 4801
Db 4758 GUUUGUUGAAUUAUUGUGUA---UUAACACAGUUGGAGCAGCGUUAUUGGAGGUU 4814
Qy 4802 CATAAACCGGATTTGATGTTGCTTTGTTTAAATTTGTTTAAACAAATTTTGTGTGAT 4861
Db 4815 CAUAGACCCGCCUCCAGGUGUUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4874
Qy 4862 AAAATTTTATTTAGAACTTTTAAATGCTGTTAGTCTCAGAGATCTGTCAAAT 4921
Db 4875 AAAGUUCUUUAAGAUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4934
Qy 4922 TAGCGATTCATTGATCTTTTCGAAACAGGATCAGATCTTCCGCAATTCATGACTAAGGT 4981
Db 4935 CAUAGAUUUAUUGACCGACCAAAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4994

Qy		140	AATGACTTTGCAGAGGCGCATGTACGATA CGGCCGTGGAGAAGATTTAA CGCCCGCGCAC	199
Db		138	AAUGAUCUACAAGCGUGCUUUUAUUAACGACACAGCGGUUGAAGAUUAAUAA CCGCUGAC	197
Qy		200	CGTAGACCRAAGGTCACTTTTCCTCAAAA CTATTAGCGAAGACCAACGCTTCTAGTCTCC	259
Db		198	CGCAGGCCAAGGUGAACUUUUCAAAGUAUAUAGCGAGGACGACGCGCUUAUUGCUACC	257
Qy		260	AACCGCTTACCGGAGTTCCAGATTACTTTTAATACTCAAAATGCCGTACACAGTTTG	319
Db		258	CGSGCGUAVCCAAGAUUUCAAAUUAUAUUAUUAACA CGCAAAUUGCCGUGCAUUCGUU	317
Qy		320	GCTGGAGGTTTTAGACGATTTAGATTTGGAATCTCTGATGCTACAGTTCCTCTATGATCG	379
Db		318	GCAGGUGGUAUUGGUAUCUUUAGAAACUUGGAAUAUCUGAUGCAAUAAUCCUUCGGAUCA	377
Qy		380	CCGACATATGATATAGTGGGAACTTTTGCAGCACATTTGTTCAAAGSCAGGATTAACGTG	439
Db		378	UUGACUUAUGACAUGCGCGGAAUUUGCAUCGCAUCUGUUA CAAGGACGAGCAUAUGUA	437
Qy		440	CATTGCTGTATGCCAATCTGGACATACGAGATATAATGAGGCACGAAGGACAAAAGNAC	499
Db		438	CACUGCGCAUGCACCAACUGGACGUGUGAGACAUC AUGCGGCACGAAGGCCAGAAAGAC	497
Qy		500	TCAATTGAGATGTATTGTCCAGATTGTCTCGTTCTTAA CAAGGTAAATTCCTGAGTTCAA	559
Db		498	AGUAUUGACAUAUACUUUUCAGGCUAGAGAGGGGGGNA AACA GUCUCCCAUUCCAA	557
Qy		560	AGGAGGCTTTTAA CAGGTTATG CAGAAGCTCCCAACGAAGTCTGCTCTTAAACTTTT	619
Db		558	AAGGAACUAUUUGACAGUAACGAGAAAAUUCUGAAGACGCGUCUGUACAUAUACUUD	617
Qy		620	CAGGATTTGCAATACATCCGCCAGAGATAGTGSTAGAGATACGCTGTGCTCTGCAAC	679
Db		618	CAGACAUGCGACAUCAGCGCAUGCAUCAUAGGCAGAGUGUAUGCCA UUGCGCUACAC	677
Qy		680	AGTTTGTATGATATTCCTGTGCATGAGTTTGGAGCTGCGTTAATATCTAAGAATATACAT	739
Db		678	AGCAUAUAUGACAUAAC CAGCCGAGUAUUGCGGGCGGCA CUUUGAGGAAAAUUGUCCAU	737
Qy		740	GTATGTTATCGACTTCCATTTTGGCAGAGCATTATTACTAGACAGACGGAGGTTACG	799
Db		738	ACGUGCUAUGCCGCUUUCACUUCUCUGAGAA CCUGCUUCUGAGAUUCAUACGUCUAAU	797
Qy		800	CTTAAATGAAATPAGCGCAACTTTTCAAAAGAGAAGGTGATGATGTTCTTTTCTTTGCT	859
Db		798	UUGACGAAAUCAACCGUGUUUUUUCGCGAUGGAGACAAGUUGACCUUUUUUUGCA	857
Qy		860	GATGAAAGTACTTTAAATTTAGTCAATAATA CAABAATATCTTGCA TTATGTAGTTAA	919
Db		858	UCAGAGAGUACUCUUAUAUUAUGCAUAGUUA UUCUAAUAUUCUUAAGUAUGUGUCAA	917
Qy		920	TCTTACTTTCTCTGCTCTAGTAGAATAAGTTTACTTTTAAGGAAATTTT TPGTCACTAGGTT	979
Db		918	ACUUAUUCGCGCCUUAUAUAGAGAGUUUAUCAUGAAGAGGUUUUAUUGACCAAGAUU	977
Qy		980	AATFACTGTTGTTTTGTAATTTTACCAAGTAGATACCTATATCTGTA CAGAGTGTAGA	1039
Db		978	AAUACCGUGUUUUGTAUUGUUUCUAGAUAUGAU CUUUUCUUUGUA CAAGGUGUGGCC	1037
Qy		1040	CAAGTAGGCTGTAGTGTATGATCAGTCTCTAGAGCGGATGGAAGACGCTTTGCTTTACAAG	1099
Db		1038	CAUPAAGUAGUAUAGUGACGAGUUUAUUAUCGCAUUGGAGAGACGCAUGGCAUUAACA	1097
Qy		1100	AAAACTTTGGCCATGTTTCAACA CTTGAAGAGCAATCTTTTAGACACACGGCTTCGTTAAC	1159
Db		1098	AAGACUCUUGCAUUGUCAA CAGCGAGAGAAUCCUUCUUGAGGAUUCAUCAUCAGUCAU	1157
Qy		1160	TTTTGTGTTCCCTAAGNTGAAGGACATGCTGATAGTACCGCTGTTTGGAGGTTCTATTACC	1219
Db		1158	UACUGUUUCCCCAAAUGAGGGAUUGUGUAUCGUA CCAUUAUUGCAUAUUCUUGGAG	1217
Qy		1220	AGCAAAAAGATGACAAGGAGTGAGGTCATTCGTTAATCGTGA CTCGTTTACAGAGTCGTT	1279

Db 2298 GUCUUGAUGUUGCAUCUAGGAAGUGUUAUUAUCAAACCAACCGCCCAAGAGUACUGAUGG 2357
Qy 2354 GGAGTTGTCCTCGATTACCAAGGGGAAATGTTTACTGCACTTCTATCTTATGAAAGAGAT 2413
Db 2358 CGUGUUGUUGAAACCAACCGAGGAAGUAUUAUGUGGCGCUUUUGAAUAUUAUGAGCAG 2417
Qy 2414 AGAATGGTGACCTGAGAGCACTGGAGGAGGCTGTATCATCTGATCAATGATATAT 2473
Db 2418 GUGUGUGGCAUGCAUGCAUUGGAAGAGAGUAGUGUGUGUGUGUGUGUGUGUGUGUUAU 2477
Qy 2474 TCTGATATTGCAAGCTCCAAATCTGAGGAAACAACTGAGAGAAGTGAACCCACGAA 2533
Db 2478 UCCGACAUUGCGAACAUCAGACUCUGCGCAGACUUGCUUGGAAACGGAGAACCGAUGUC 2537
Qy 2534 CTAATCTGCAAGATGGTACTTGTGAGTGGGCTGCTGTGTTGGAAAGTACAAAGAGAT 2593
Db 2538 AGUAGCGCAAGAGUUGUUGUGGACGAGUUGCGGUGUGGAGAAACCAAGAAAUU 2597
Qy 2594 TTTGAAAGATTGATCTTGATGAGATTGATCTTGTTCTGGAACCAAGCTGCTGCT 2653
Db 2598 CUUUCGAGGUUAUUAUGAAGAGUAUUAUUAUGUACUUGGAAACGAGCGCGGAA 2657
Qy 2654 ATGATCAGAAGAGGCTAAATCATCTGAGCTGATAAGAGCCACAAATGCAATGTGAGA 2713
Db 2658 AUGAUCAGAGAGUGCGNAUUCUAGGGAUUAUUGUGGCCACAGAGGACAAACGUUAAA 2717
Qy 2714 ACGGTAGATTCTACTTAATGATCCAAACCCGCGATCACA-----CAAGAGGCTT 2764
Db 2718 ACCGUUGAUUCUUAUGAUGAAUUUUGGAAAGCACACAGGUGUCAGUUAAGAGGUUA 2777
Qy 2765 TTTATTGATGAAGGTTGATGCTGCACACCGGTTGTGTTAACTTCTGGTGCTTATCTCT 2824
Db 2778 UUAUGAAGAGGUGAUGUGUAUCUGUGUGUUAUUAUUCUGUGCGCAUGUCA 2837
Qy 2825 GGTTCGACATCGCATACATTTACGAGATACACAGCAGATCTCTTTCATTAACAGAGTT 2884
Db 2838 UUGUGGAAAUUGCAUUAUUGUUAACGAGACACACAGCAGAUUCCCAUUAUUAUUAU 2897
Qy 2885 CAGAAATTCCTGATCCCAACATTTGAGAGCTGCAAGTGGATGAAAGTTGAGATGAGG 2944
Db 2898 UCAGAUUCCGUAUCCCGCCCAUUAUUGCAAUUGAAGUUGACGAGGUGGAGACACGC 2957
Qy 2945 AGGACCACTGAGATGCGCAGGTGATGTGAATTTTCTTCAATCGAATGTAAGGA 3004
Db 2958 AGAACUACUCCUGUGUCCAGCGCAUGUACAACAUUAUCUGAAACAGGAGAUUAUGAGGC 3017
Qy 3005 GCGGTGACAAACCACTTCAACTGTACACGATCGGTCTCATCTGAGATGATAGGCGGTAA 3064
Db 3018 UUUUGAUGAGCAUUCUUGGUUAAAGUUGUUAAGUUGUUGCGAGGAGUGGUGCGGAGCC 3077
Qy 3065 GGAGTACTAAACAGTGTTCAAACCACTTAAAGGGAATAATTTGTAATTTCACTCAGGCT 3124
Db 3078 GCGUGAUCAUCCGUAUCCGAAACCCUUGCAUGGCAAGAUCCUGACUUAUUAUUAUUAU 3137
Qy 3125 GATAAATTTGATTAGAGAGAGGCTATAGAAATGTGAACACGGTTTCATGAGATCCAA 3184
Db 3138 GAUAAAGAGCUCUCUUAAGAGGUAUUAUGAUGUUCACUGUGCAUGAAGUACA 3197
Qy 3185 GGAGAAACCTTTGAAGATGTGCTGTGTGATGAGCGCACTTCCACTGACTCTGATT 3244
Db 3198 GCGAGACACUACUGAUGUUUUAUUAAGUUAAGUUAAGUUAAGUUAAGUUAAGUUAAGU 3257
Qy 3245 TCCAAATCTTCCCGCATGTTCTAGTCTGCTGATAGACACACAAAGAGCTTCAAAAT 3304
Db 3258 GCAGGAGACAGCCCAUGUUUGGUGCAUUGUUAAGGCAACCCUGUUGCUCUUAUUAUUA 3317
Qy 3305 TACACCGTAGTGTAGATCTTTAGTACAGATAATAGTGTGCTTCTTTTAACTCC 3364
Db 3318 UACACUGUUGUUAUGGAUCCUUAUAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3377
Qy 3365 TTCTTTTAAAGATGATATGTAAGACGAGTGTAGATGATGATGATGATGATGATGATGAT 3424
Db 3378 UACUUGUUAUGAUGUUAUAGUGUAGUAGGAAACAAUUAAGCAUUAUUAUUAUUAUUAUUA 3437

Qy 3425 GTGTTCAAAGTCTAATCTCTTTGTGGCAACACTAAATCAGGAGACTTTCCAGATCTA 3484
Db 3438 GUGUUAAGGUGUCCAAUUCUUUUUGCAGCGCCAAAGACUGGUGAUUUUCUGAUUG 3497
Qy 3485 CAGTTCTATTACGATGATGCTCCCTGGTAAATAGTACTATATCTTAAACAAGTATGCT 3544
Db 3498 CAGUUUAUUAUGAUGUUGUCUCCAGGCAACAGCACCAUGAUAUUAUUAUUAUUAUGUC 3557
Qy 3545 GTTACCATGAGTTACGTGATAATAGTCTTAATGTGAAGGATTTGTTCTTGAATTTTCC 3604
Db 3558 GUUACCAUGAGUGAUGACAUUUAUGAUGAUGAUGAUGAUGAUGAUGAUGAUGAUGAUGUC 3617
Qy 3605 AAAAGTATTCGATGCGCAAGAGGTAACCATGTCTTAGAGCCAGTTTTCGTTACCGC 3664
Db 3618 AAGUCUGUGUGCGCCUUAAGGAUCAAUCAAACCAUUAUUAUUAUGUACGAAACGCGC 3677
Qy 3665 GCGGAACCGCAAGGCTGCGAGGACTACTCAGAAATCTGTTGCATGTTTAAAGAAAT 3724
Db 3678 GCAGAAUUGCCACGCGCAGACUGGACUUAUGAUAUUAUGGCGGAUGAUAUUAAGGAAC 3737
Qy 3725 TTCAACGCAACAGACCTGACGCGGACGATTGACATTTGAGAGCACCGCATCTGTTGATGA 3784
Db 3738 UUAACGCAACCGAGUGUGUGCAUUAUGAUAUUAUGAUAUUAUGAUAUUAUGAUAUUAUG 3797
Qy 3785 GATAAGTTTGTGATGCTATTTTAAAGAAAGAAATAACAAAAAATATTTGCTGGA 3844
Db 3798 GAUAAUUUUUUAUGAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3857
Qy 3845 GTGATGACGAAGGATTCATGATGAGATGTTGAAAGAACAGAAAGAGTACTATTGGAC 3904
Db 3858 UUGUUCAGAGAGAGUCUCUUAUAGAUUGUUAAGAAAGAGAGAACAGUUAACAUAGGC 3917
Qy 3905 GACTTGGCTAACTACAATTTTACAGATCTGCGGCGCATCTGATCAGTCAAGACATGATC 3964
Db 3918 CAGCUGCAGAUUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3977
Qy 3965 AAGGCTCAACCAAAACAGAAATTTGACCTTTCAATTCAGAAATGAATACCTGCTCTGCA 4024
Db 3978 AAAGCACAACCAAGCAAAAUUAUGACACUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4037
Qy 4025 ACAATTTGCTACATTCGAGCAGATCAACGGTATTTTGGCGGTTTCT---CAGAGCTT 4081
Db 4038 ACGAUGUGUACCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4097
Qy 4082 ACAAGTTTGTCTCGAGGCAATTTGATTTCTAAGAGTTTCTTTTCTTTCTTCTAGGAAACT 4141
Db 4098 ACUAGGCAUUAUUAUGGACAGUUGAUAUUAUGGACAGAUUUUUUUUUUUUUAUUAUUA 4157
Qy 4142 CCAGAACAGATTCAGAAATTTTCTCGGATCTCGACTCGCACTGCTTCTCTATGATGTTTA 4201
Db 4158 CCAGCGCAGAUUGAGGAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4217
Qy 4202 GAATCTGATATTTCTAAGTATGATAGTCAAGAACAGATTTTCTTGTGCTGTAGATGAT 4261
Db 4218 GAGCUGAUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 4277
Qy 4262 GAAATATGAAAGATTTGGGTCTCAATGAGTTTGGCCGAGGTTGCGAAGGTAAGGAGGAC 4321
Db 4278 GAGAUUUGCGAAGAUUGGUGUUUAAGAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4337
Qy 4322 AGGAAAAACAATTTGAAGGATTAATGCTGGAATCAAGACATGCTGTGTTGATCAAAAG 4381
Db 4338 AGAAAGACCAACCCUUAAGGAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 4397
Qy 4382 AAAAGCGGTGATGATCTATTTTCTATCGGCAATCTGTTTAAATAGAGCTTGTGTTGGT 4441
Db 4398 AAGAGCGGAGCAGCAGCAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 4457
Qy 4442 TCAATGTTTACGATGGAAGGTCATAAAGGTCCTTTTGTGAGAGCGATTCGCTTTTG 4501
Db 4458 UCGAUGUCUCCGAGGAGAAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 4517

Db 1818 UUGGAAGUUGACCCAAUAGCGCAGCAAGGUUAUAGUCGGGUCAUGAGCAUAGAGAGC 1877
Qy 1880 GGTTTAACTCTTACTTTTATGATTAAGCAACCGAGGAGATGTGGCTTAAAGCTCT-----T 1933
Db 1878 GGUCGACUCUCACAUUUAAGACGACUACUGAGGGGAUUGUUGUCUAGCUUUAACAGGAU 1937
Qy 1934 AAAAGCAGCGCTCTGAGCGCTGTGTATGTCTTTAAACGACATCCGAAGAGGTGAACGTA 1993
Db 1938 CAAGAGAGGCUUUCAGAGGUGCUUUGUAGUUAUCCUACAAGAGAGUUGAAGAACCGUCC 1997
Qy 1994 AATAAATTTTCTATTGCTGAGAAAGGAGATTGCTGTGTGTCGAGAAAGTCAATGTTTG 2053
Db 1998 AUGAAGGUGUUGAUGCCAGAGAGAGUUAUCAAUUGAGUGGUCUGGAGAUCAUCCG 2057
Qy 2054 ACGAATGCTAACTTAGACACACAGAGTTGGAGTCCCTCAACGATTTCCATAAGCTTGC 2113
Db 2058 GAGUGUCUUAUUCURAGAACAGAGAGUAGAGUCUUUAGAGCAGUUAUUAUAGCAACG 2117
Qy 2114 GTGGATGTGTGATTACAAAGCAATGGCATCGGTGTCTACATCGGCTCACTCAAGTT 2173
Db 2118 GCAGAUUCGUUAUUCGUUAGCAGAGAGCUCGAUUGUUAUACCGGUCUUGAUUAAGUU 2177
Qy 2174 CAACAAATGAAGAACTATGTGACAGATTGGCAGCTTGGTGTGTCGCCCACTGTATCAAT 2233
Db 2178 CAGCAAAUGAAAAACUUUAUCGAUAGCCUGGUAUCAUUAUCUGUCUGCGUGUGGAU 2237
Qy 2234 CTATGCAAGTCACTAAGAGATGAAGTCGGTATGATTCTGATTCAGGGAGAAAGTTGTT 2293
Db 2238 CUCGUCAGAUCCUCAAAGUAUACAGCUCUUAUUGACCUUGAAACCGUCAAAGUUGGA 2297
Qy 2294 GTTTGGGATGTCACTTTGAAAGTGGCTCTCAACCTCGCGCCAAAGGTCATTCTATCG 2353
Db 2298 GUCUUGAUGUUGCAUCUAGAGAGUUGUUAUCAAACCAACCGCCAGAGUUAUGCAUGG 2357
Qy 2354 GGAGTTGCTCTGGATTAAAGGGGAAATGTTTATCTGCACTCTCTATCTTATGAAGAGAT 2413
Db 2358 GGUGUUGUUGAAACCCACGCGAGGAUUAUUAUGUGGCGCUUUGGAUUAUUAUGAGCAG 2417
Qy 2414 AGAATGCTCACTGAGAGCACTGGAGAGGGTGTGTATCTATCTGATCAATGTTATAT 2473
Db 2418 GGUGUGGCAUUGCAUGCAUUGAUGAGAGUAGUGCUGCAGCUCUGAGUCUGUUGUUAU 2477
Qy 2474 TCTGATTAATGCAAGCTCCAAATCTGAGGAAACAAATGAGAGACGGTGAACCCACGAA 2533
Db 2478 UCCGCAUGCGGAAACUCAGAAACUCUGCGCAGACUUGUUGGAAACGAGAAACCGCAUG 2537
Qy 2534 CCTACTGCAAGATGTTACTTGTGATGGGGTCCCTGTTGTGGAAAGTACAAAGAGAT 2593
Db 2538 AGUAGCGCAAGGUGUUGUUGGACCGAGUUCGCGGUGUGGGAACCAAGAAAUU 2597
Qy 2594 TTTGAAGATTGATCTTCATGAGGATTTGATCTTGTGTTCTTGAACCAAGCTGCTGCT 2653
Db 2598 CUUUCGAGGUUAUUUGAUGAAGAUUAUUUUUAGUACUUGGAGACGACCGCGAA 2657
Qy 2654 ATGATCAGAGAGGGCTTAATCTATCTGACTGATAAGAGCCCAATGACAAATGTGAGA 2713
Db 2658 AUGAUCAGAGAAGCUGCGGAUUCUUCAGGAUUAUUGUGGCCACGAGGACAAACGUAAA 2717
Qy 2714 ACGGTAGATTCACTTCTAATGATCCAAACCGCGATACA-----CAAGAGCTT 2764
Db 2718 ACCGUUGAUUCUUAUGAUGAAUUUUGGAAAGCACACGUCUGUUAUUAAGAGGUUA 2777
Qy 2765 TTTATTGATGAAGGGTTGATCTGCACACCGGTTGTGTTAACTTCTCGTGTCTTATCTCT 2824
Db 2778 UCUUAUGAUGAGGGUUGAUGUUAUUAUGUUGUUAUUAUUUUGUGGCAUGUCA 2837
Qy 2825 GGTTCGACATCGCATACATTTACGAGATACACAGCAGATCTCTTTCAATTAACAGATT 2884
Db 2838 UUGUGCGAAAUUGAUUUAUCGAGACACACAGCAGAUUUAUUAUUAUUAUUAUUAUUA 2897
Qy 2885 CAGAAATTTCCGATCCCAACATTTTGAAGAGCTGCAAGTGAATGATGAGG 2944

Db 2898 UCAGAUUCCGUAACCCCGCCCAUUUUGCCAAAUUGGAUUGUAGGAGGUGGACACGC 2957
Qy 2945 AGGACCACTGAGATCCAGGTGATGTGAATTTTTTCTTACAAATCGAAGTACGAAGGA 3004
Db 2958 AGAACUACUCCUUGUCCAGCCGAGUACACAUUAUCUGACAGAGAUUAGAGGC 3017
Qy 3005 CGGTGACAACTTCAACTGTATCAACGATCGGTCTCATCTGAGATGATAGGCGGTAA 3064
Db 3018 UUGUCAUGAGCACUUCUGUUAUAAAGUUCUGCGAGAGAGUGGCGGAGCC 3077
Qy 3065 GGAGTACTAATACAGTGTTCACAAACCACTAAAGGAAAAATTTGTAATCTTTCACCTCAGCT 3124
Db 3078 GCCGUGAUCAUCCGUAUCCAAACCCUUGCAUGGCAAGAUCCUGAUUUUUAUCCCAUUG 3137
Qy 3125 GATAAATTTAGTTAGAGGAGAGGCTTATAGAAATGTGAAACACCGTTTCATGAGATCCA 3184
Db 3138 GAUAAAGAGCUCUGCUUUAAGAGGUUAUUCAGAUUGUACACUGUGCAUAGAGCAA 3197
Qy 3185 GGAGAAACCTTTGAGATGTGTCTGTGTGATTTGACGGCAACTCCACTGACTCTGATT 3244
Db 3198 GCGGAGACAUAUCUGUGUUAUACUAUAGGUUAACCCCUACACCAAGUCUCCAUCAU 3257
Qy 3245 TCCAAGTCTTCCCGCATGTTCTAGTCTGCTAGCTAGACACACAAAGAGCTTCAAAAT 3304
Db 3258 GCAGAGACGCCACAUUUUUUGUGCAUUGUACAGGACACACUGUUCGCUCAAGUAC 3317
Qy 3305 TACACCGTAGTTAGATCTCTTTAGTACAGATAATTAGTGTGTTGTTCTTTTAACTGCC 3364
Db 3318 UACACUGUUAUGGAUCCUUAAGUUAUUAUAGAUUAUUAAGAGAUCAUUAUUGCUG 3377
Qy 3365 TTCTTTTAGAAATGTATGTGTAGAGCAGGTAGTAGTAATTAACATTCACATGATGCA 3424
Db 3378 UACUUGUUAUUAUUAAGUGCAUAGCAGGAAACAAUAGCAUUAUUAUUAUUAUUGC 3437
Qy 3425 GTGTTTAAAGTCTAATCTCTTTGTGCAACACCTTAATCAGGAGACTTTCAGATCTA 3484
Db 3438 GUGUUAAGGUGUCCAUUUUUUGUGCAGCGCCAAAGACUGGUAUUAUUAUUAUUG 3497
Qy 3485 CAGTTCTATTACGATGTATGCTCCCTGGTAAATAGTACTATATCTTAAACAATGATGCT 3544
Db 3498 CAGUUUAUUAUUAUUAAGUGUCUCCAGGCAACAGCACCAUGAUUAUUAUUAUUG 3557
Qy 3545 GTTACCATGAGTTACGTGATAATAGTCTTAAATGTGAAGGATTTGTTCTTGTATTTCC 3604
Db 3558 GUUACCAUGAGUGUACUGCAUUAUUAUUAUGUCAAAGAUUGCAUUAUUAUUAUUG 3617
Qy 3605 AAAAGTATTCGATGCCAAAGAGGTGAAACCATGCTCTAGAGCCAGTTTGTGTTACCGG 3664
Db 3618 AAGUCUGUUGCGCCUUAAGGAUCAAACCAUUAUUAUUAUUAUUAUUAUUAUUG 3677
Qy 3665 GCGGAACCCGCAAGGCTGCAGGACTACTCGAAATCTGGTTGCAATGATTAAAGAAAT 3724
Db 3678 GCAGAAAUUGCCACGCGAGACUGGCUUAUUGGAAAUUAUUGGCGCAUUAUUAUUAU 3737
Qy 3725 TTCAACGACCAACGACTCGCGGGAGATTGACATTTGAGAGCCCGCATCTGTTCTAGTA 3784
Db 3738 UUAUAAACGACCGAGUGUGCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3797
Qy 3785 GATAAGTTTGTGATGCTATTTTATTAAGAAAAAATAACAATAATAATTTCTGGA 3844
Db 3798 GAUAAAGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3857
Qy 3845 GTGATGACGAAGGATTCATATGATGATGTTGGAAAAACAGGAAGAAGTACTATTGGAC 3904
Db 3858 UUGUUCAGUAGAGAGUCUCUAAUAGAUUGGUUAAGAAAGAGAGAAACAGUAAACA 3917
Qy 3905 GACTTGGCTTAATACTAATTTTACAGATCTCGCGGCATCTGATCAGTACAGACATGATC 3964
Db 3918 CAGCUCGAGAUUUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAU 3977
Qy 3965 AAGGCTCAACCAAAAAACAGAAATTTGACCTTTTCAATTCAGAAATGAATACCTGCTCTGCA 4024
Db 3978 AAGACCAACCCCAAGCAAAAAUUGGACACUUAUUAUUAUUAUUAUUAUUAUUAU 4037

QY 1340 GAGTCTATAAGATCCCGGTGATATCAATGGTGTGTTACTGTAGTCTGAATGGATGTA 1399
Db 1338 GAUUGAUUUGAUCGAGGUAUAUAUAACGUGUGACAGCGAGGUGCCGAUUGGUG 1397
QY 1400 GATTAAGCAATTTCTTCAACCCCTGTGTCAATGACTTTCTTCTTGCGAGCTAAGCTGGCTGCG 1459
Db 1398 GACAAAUUUGUUAUAAUCCUUGUCCAGAGUUUAUACCUUGCAUAUAAGCUUGCCGU 1457
QY 1460 CTTCAAGACGATATAGTAATGGGAAGTTTGGTCTGTTGGATAGACAACATTTCTGAATTT 1519
Db 1458 CUAAGAGGAGCAUAUAUAUAAGCAAGUUUAUAGUCUGCGTUGGAAACGGUGUGCCAGCAU 1517
QY 1520 ATTTGGGATGGTGGGCAATTTTTCGAAACGTTTTCCTCCACTATCAAGAGAGATG 1579
Db 1518 GUGUGGAGUAGAUUUCGUGCGUGGGAACGCAUUCUUGGAAAGAGAGGCGUC 1577
QY 1580 GTGACGAGCAAAATTTCTGGATTAAGTGAGATGCTCTGAAGATCAAGATCCAGATCTG 1639
Db 1578 UUGAACAGGAACUAUAUCAGAGUGCGAGCGACGCAUUAUAGAUACAGGUGGCTGUAUA 1637
QY 1640 TATGTCAATGGAAGAAGAGTTCTGTAGTCTGAATACAAAGTCTGAGAGTTACCGCAT 1699
Db 1638 UAUGUGACCUUCCACGACAGAUUAGUCAGUGAUAAGGCGCUCUGUGCAUUGCCUGCG 1697
QY 1700 CTAGATATCAAGAGGACTTAGAAGAGCTGAGCAAAATGTAGCGCGCTTATCAGAAATTA 1759
Db 1698 CUUGACUAUAGGAAGAAGUAGAAACGGAAGUGUGAUAUUGGCAUUCUUCAGAGUUA 1757
QY 1760 TCTATCTTAAAGGGTCTGATAATTTTCGATATCGGAAGTTTCAAGAGACATGTGCAAGGCT 1819
Db 1758 UCGGUGUAUAGGAGUCUGACAAUUCUUAUUGUUAUUGUUAUUCUCCAGUUGCCNAUCU 1817
QY 1820 TTAGATGTTAGTCTGTGTGACGACGAGTAATGTTGTCAGTGGCGGAGAAATAGAAC 1879
Db 1818 UUGGAAGUUGACCCAAUGACGCGAGGAGUUAUAGUCGCGUGCAUAGAGCAUAGAGAGC 1877
QY 1880 GGTTTAACTTACTTTTGTATAAGCAACCGAGGAGATGTGGCTTAAGGCTCT-----T 1933
Db 1878 GGUCGACUCUACAUUUGAACGACCUACUAGAGCGGAUUGUGCGCUUUAACAGGAU 1937
QY 1934 AAAACACCGGCTGTAGCGCGTGTAATGTTGAAACCGACATCCGAAGAGGTGAACGTA 1993
Db 1938 CAAGAGAGGCUUUGAAGAGGUGCUUUGUUAUUAUCCUACAGAGAUAUAGAACCGGUCC 1997
QY 1994 AATAAATTTCTTAATGTCGAGAAAGGAGATGTCCTGTGTGTCGAGAAGTCAATGTTG 2053
Db 1998 AUGAAGGUGUUGAUGCGCCAGAGGAGUUAUAAUUAUUGUUGUUGUGGAGAUCAUCG 2057
QY 2054 AGAATGTAATTAAGACACAGAGGTTGAGTCCCTCAAGATTTCCATAAGGCTTGC 2113
Db 2058 GAGUCGUCUAUUCUAAGAACGAGAGAUAGAGUCUUAUAGAGCAUUAUAGGCAACG 2117
QY 2114 GTGGATAGTGTATTACAAAGCAATGGATGCTGGTGTCTACCTGGCTCACTCAAGTT 2173
Db 2118 GAGAUUCGUUAUUCGUAAGCAAGAGAGCUGCUUUGUGUAUACCGGUGCCGAUUAAGUU 2177
QY 2174 CAACAAATGAAGAACTATGTGACAGATTTGGCAGCTTGGTGTGTCGCGCACTGTATCAAT 2233
Db 2178 CAGCAAAUAGAAAACUUUAUCGAUAGCCUGUAGCAUACUAUUGUGCGGUGUGRAU 2237
QY 2234 CTATCAAGTCACTAAAGATGAAGTCCGGTATGATTTCTGATTCAGGAGAAAGTTGGT 2293
Db 2238 CUCGCAAGAUCCUCAAAGAUACAGCUGCUUAUAGACCUUGAAACCCGUCAAAAGUUGGA 2297
QY 2294 GTTTGGGATGTCACTTTGAAAGTGGCTCTCAACCTGGCGGCAAGAGTCAATCATGG 2353
Db 2298 GUCUGGAUGUUGCAUCUUGGAAGUGGUUAUACAAACCAACGCGCCAGAGUUAUGG 2357
QY 2354 GGAGTGTCTGGAATTACAAAGGGGAAATGTTTACTGTGCACTTCTATCTTATGAAGAGAT 2413
Db 2358 GUGUUGUUGAAGAACCCACGCGAGGAGUUAUAGUGGCGCUUUGGAUUAUAGAGCAG 2417
QY 2414 AGAATGGTGACTGAGAGCGACTGGAGGAGGGGTGTATCATCTGATACATATGATATAT 2473

Db 2418 GGUGUGUGAGCAUCGCGAUGAUUGGAAGAGAGUACGUCAGCUCUGAGUCUGUUGUUUAU 2477
QY 2474 TCTGATATTGCAAGCTCCAAAATCTGAGGAAAAACAATGAGAGACGGTGAACCCCAACAA 2533
Db 2478 UCCGACAUUGGCGAAACUCAGAAACUCUGCGCAGACUGUCUUCGAAAACGGAGAACCGCAUGUC 2537
QY 2534 CCTACTCMAAGATGGTACTTGTGTGATGGGTGCTGCTGTGTGTGGAAGTACAAAGAGAT 2593
Db 2538 AGUAGCCAAAAGGUGUUCUUGUGGACGGAGUUCGCGGUCUGGGAAAAACCAAGAAAAUU 2597
QY 2594 TTTCAAGAGATTGATCTTGATGAGATTGATCTTGTTGTTCTCGAAAAACAAGCTGTCTCT 2653
Db 2598 CUUCCAGGGUUAUUAUUGAAGAGAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2657
QY 2654 ATGATCAAGAAAGGGCTTAATCTCTGACTGATAGAGCAACAATGGAACAATGTGTAGA 2713
Db 2658 AUGAUCAGAGACGUGCGAAUUCUCAGGGAUAUUAUUGGCGCACGAGAGCAACGUAUAA 2717
QY 2714 ACGGTAGATTCACTTTAATGCAATCCAAAACCCGCGATCACA-----CAAGAGGCTT 2764
Db 2718 ACCUGAUUUCUUAUGAUGAAUUUUUGGAAAAAGCACACGUCUGACAUUAAGAGGUUA 2777
QY 2765 TTTATTGATGAAGGTTGATGCTGCACACCGGTTGTGTTAACTTCTCTGCTGCTTATCTCT 2824
Db 2778 UUCAUGAAGAGGUGAUGUUGCAUUCUGUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2837
QY 2825 GGTTCGACATCGCATATTTACGGAGATACACAGCAGAGATTTCTTTTCATTAACAGAGTT 2884
Db 2838 UUGUGGAAAUUGCAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 2897
QY 2885 CAGAAATTTCCGTTATCCAAAACATTTTGAGAGCTGCAAGTGGATGAAGTTGAGATGAGG 2944
Db 2898 UCAGGAUUCGUAACCCGCGCCAUUUUGCAAUUUGGAAGUUGACGAGGUGGAGACGCG 2957
QY 2945 AGGACACATGAGATGCCAGGTGATGTGAATTTTCTTCAATCAATCAAGTACGAAGCA 3004
Db 2958 AGAUAUAUCUCCUGUGUCCAGCGCAUGUACAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3017
QY 3005 GCGGTGACCAACCTTCAACTGTACAACGATCGGTCTCATCTGAGATGATAGGCGGTAAAG 3064
Db 3018 UUGUCAUGAGCACUUCUUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3077
QY 3065 GGAGTACTAAACAGTGTTCAAAACCACTAAAGGAAAAATGTAACTTTCACTCAGGCT 3124
Db 3078 GCCGUGAUCUACGCAUCCCAAAACCCUUGCAUGGCAAGACUCCUGAUUUUACCCAAUCG 3137
QY 3125 GATAAATTTGAGTTAGAGGAGAGGCTTATAAGAAATGTGAACACGTTTCAATGAGATCCAA 3184
Db 3138 GAUAAAGAGCUCUGCUUUAAGAGGUAUUAUGAUUUAUUAUGAUUAUUAUGAUUAUUAUGCAA 3197
QY 3185 GGAGAAACCTTTGAAGATGTGCTGCTGTCAGATTGAGCGGCAACTCCACTGACTCTGAT 3244
Db 3198 GCGAGACAUAUCUGAUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3257
QY 3245 TCCAAGTCTTCCCGCATGTTCTAGTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 3304
Db 3258 GCAGGAGACGCCACAUUUGUGCAUUGUCAUUGCAAGGACACCCUGUUCGUCUUAAGUAC 3317
QY 3305 TACACCGTAGTGTAGATCTTTTAGTACAGATAATTAGTGTGTTGTTGTTGTTGTTGTTG 3364
Db 3318 UACACUGUUAUUGGAUCCUUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3377
QY 3365 TTCCTTTTAGAAATGTATGTTAGAGACAGTGTAGTAGTAATGAGTAATACAGATGGATGCA 3424
Db 3378 UACUUGUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUAUUA 3437
QY 3425 GTGTTCAAGGTCAATATCTCTTTGTGGCAACACTAAATCAGGAGACTTTCCAGATCTA 3484
Db 3438 GUGUCAAGGUCCAAUUCUUUUGUGCAGCGCGCAAGAGACUGGUGAUUAUUAUUAUUA 3497
QY 3485 CAGTTCTTATAGGATGATGCTCCCTGGTAAATAGTACTATCTTAACTAAGATGATGCT 3544

Db 3018 UUUGUAGAGCAUUCUUCGUUUAUAAAGUUCUUCGAGGAGUUGUGCGGGAGCC 3077
Qy 3065 GGAGTACTAAACAGTGTTCACAAACCACTAAAGAGGAAATTTGAATCTTTCACCTCAGGCT 3124
Db 3078 GCCGUGAUCAUCCGAUCCAAACCCUUGCAUGGCAAGAUCCUGAUUACCAUUG 3137
Qy 3125 GATAAATTTAGTTAGAGGAGAGGGCTATAAGATGTGAACACCGTTCATGAGATCCAA 3184
Db 3138 GAUAAAGAGCUCUGUCUUUCAAGAGGUAUUCAGAUUGUUCACUGUGCAUGAAGUGCAA 3197
Qy 3185 GGAGAAACCTTTGAAGATGTGCGTGTGAGTGTGAGGCACTTCCACCTGACTCTGAT 3244
Db 3198 GCGGAGACAUACUGAUGUUCACUAGUAGGUUAUACCCUACACAGUUCUACAU 3257
Qy 3245 TCCAAGTCTTCCCGCATGTTCTAGTCTGCTAGTAGACACACAAAGAGCTTCAAATAT 3304
Db 3258 GCAGGAGACAGCCCAUGUUCGUGCAUUGUCAAGGCCACACUUGUCUCAAGUAC 3317
Qy 3305 TACACCGTAGTTAGATCCTTTAGTACAGATATTTAGTGTGTTCTTCTTAAAGCTCC 3364
Db 3318 UACACUGUUGUUAUGGAUCCUUUAGUUAUUAUAGAGAUUCUAGAGAAACUUAUGCUG 3377
Qy 3365 TTCCCTTTAGAAATGTAATGTAGAACAGAGTAGTAGATGAGCAATACAGATGGATGCA 3424
Db 3378 UACUUGUUAUUAUGUUAUGUUGUAGUAGGACACAAUAGCAUUAUACAUUAGUACUG 3437
Qy 3425 GTGTTCAAAGGTCTAATCTCTTTGTGGCAACACCTAAATCAGGAGACTTTCCAGATCTA 3484
Db 3438 GUGUUAAGGUGUCCAAUUCUUUUGUGCGGCCAAAGACUGGUGAUUUCUGAUUG 3497
Qy 3485 CAGTCTATTAGGATGATGCTCCTCGTGTATAGTACTATCTTAAACAGATGATGCT 3544
Db 3498 CAGUUAUUAUGAUUAGUGUCCCGCAGGCAACAGCACCAUGAUGAAUUAUUGAUGCU 3557
Qy 3545 GTTACCATGAGTACGTGATAATAGTCTTATGTGAAGGATGTGTTCTTGATTTTCC 3604
Db 3558 GUUACCAUGAGUUGACUGACAUUUCUUAUGUUAUGUUAUGUUAUGUUAUGUUAUGU 3617
Qy 3605 AAAAGTATTCGATCCCAAGAGAGTGAAACCATGTCTAGAGCCAGTTTTCGTCACGCG 3664
Db 3618 AAGUCUGUGUGCGCCUAGGAUCAAUCAAACCAUUAUACCUUAGUUAAGCAAGCGCG 3677
Qy 3665 GCGGAACCGCCNAGGGCTGAGGACTACTCGAAATCTGTTGGCAATGATTAAGAAT 3724
Db 3678 GCAGAAUUGCCACGCGCAGACUGGACUUAUUGGAAAUUUAUGGCGCAUUAUAAAGGAAC 3737
Qy 3725 TTCAAGCACCAGACTCTAGCGGGACGATTGACATTGAGACACGCGCATCTGTTGTAGTA 3784
Db 3738 UUAACGCCACCGAGUGUGUGCAUUGAUUUAUGUUAUUGUUAUUGUUAUUGUUAUGUUA 3797
Qy 3785 GATAAGTTTTTGTAGCTATTTTTATTAAGAAAGAAATAACAAATAATTTGCTGGA 3844
Db 3798 GAUAAAGUUUUUUAUUAUUGUUAUUAAGAAAGAAAGAAACCAUUAUUAUUAUUAU 3857
Qy 3845 GTGATGACGAGGATTAATGATGAGATGTTGGAAACACAGAAAGAGTACTATTGGAC 3904
Db 3858 UUGUUCAGUAGAGUGUCUCAUUAUGUUGUUAAGAAAGACAGAAACAGGUAACAAUAGGC 3917
Qy 3905 GACTTGGCTTAACATACTTTTACAGATCTCGCGGCATCTGATCAGTACAGCACATGATC 3964
Db 3918 CAGCUGCGAGAUUUUGAUUUUGAUUUUGCCAGCAUGUUAUGUUAUGUUAUGUUAUGU 3977
Qy 3965 AAGGCTCAACAAAAACAGAAATTTGACCTTTCAAATTCAGAAATGAATACCTGCTCGAA 4024
Db 3978 AAAGCACAACCCAGCAAAAAUUGGACACUUAUCCAAACGAGUACCCGCGUUUGCAG 4037
Qy 4025 ACAATTGCTACATTCGAGCAGATCAACGGTATTTTGGCCGGTTTCT---CAGAGCTT 4081
Db 4038 ACGAUUGUGUACCAUUCAAAAAAGAUCAUGCAAUUUAUUGGCCGUGUUAUGAGUCU 4097
Qy 4082 ACAAGGTTGCTGCTCGAGCAATTTCTTCTAGAGATTTCTTTTCTTACTAGGAAACT 4141
Db 4098 ACTUGGCAUUAUCUGGACAGUGUUAUUCGAGCAUUAUUGUUAUUAUUAUUAUUAUUA 4157

Qy 4142 CCAGAACAGATTCAAGAAATTTTCTCGATCTCGACTCGCACGTTCTCTATGATGTGTTA 4201
Db 4158 CCAGCGCAGAUUGAGGAUUCUUCGGAUUCUCCAGACAGUCCGCGGAUGUUGUUG 4217
Qy 4202 GAATCTGATATTTCTTAAGTATGATAAGTCAAGAAACGAGTTTCTATTTGCTGTAGAGTAT 4261
Db 4218 GAGCUGAUUAUCAAUAUACGCAAAUUCUAGAAUUAUCCACUGUGCAGUAGAUAC 4277
Qy 4262 GAAATATGAAAAAGATTGGGTCTCAATGAGTTTTTTGGCCGAAGTGTGGAACAAGGGCAC 4321
Db 4278 GAGAUUCGCGCAUGUUGGUUUAAGAAGAUUUCUUGGAGAAAGUUGGAAACAAGGGCAU 4337
Qy 4322 AGGAAACAATTTGAAGGATTACATTTGCTGGAATCAAGACATCTCTGCTGATCAAGG 4381
Db 4338 AGAAAGACCAACCCUACAGGAUUAUACCGCAGGAUUAUAAACUUGCAUCUGGUUAUCAAAGA 4397
Qy 4382 AAAAGCGGTGATGTACTTTTCATCGGCAATACTGTATATAAGACAGCTTGTCTGGGT 4441
Db 4398 AAGAGCGGGACGUCACGACUUAUUGAAACACUGAUCAUUGCUGCAUUGUUGGCC 4457
Qy 4442 TCAATGTTACGATGGAAGGTCTATAAAGGTCTTTTGTGGAGAGATTCGGTTTTG 4501
Db 4458 UCGAUGCUCUGGAGGAGAAUUAUACAAAGGAGCCUUUUGCGGUGACGAUAGUCUG 4517
Qy 4502 TATTTTCAAAGGTTGGATTTCCCTGACATTCAGTCACTGTCTATCTCATGTGGAAT 4561
Db 4518 UACUUCCAAAGGUGUGAUGUUCGGAUGUGCAACACUCCGCGAAUUCUUAUGGAAU 4577
Qy 4562 TTTGAGGCCAACTGTATAGAAAGAGGTACCGTTTACTTTTGTGTAGATACATCATACAC 4621
Db 4578 UUGAGCAAAACUUGUUAUAAACAGAUUGAUUUUUGCGGAAGAUUAGUUAUACAU 4637
Qy 4622 CATGATAAGGAGCAATAGTGTATTATGATCCTTTGAAGTTGATCTCAAACTTGGGGCA 4681
Db 4638 CACGACAGAGCAUUGUGUAUUAUGAUCUCCUAAAGUUGAUCUCGAAACUUGUGUCU 4697
Qy 4682 AAAATATCAAGGATTTGATCATTAGAGAGTTAAGGTTGTTTGTGCGATGTTGCT 4741
Db 4698 AAACACAUCAAGGAUUGGAAACACUUGAGAGAGUUCAGAAAGGUCUUCUUAUGAUGUUGCU 4757
Qy 4742 TGTTGCTCGGAAACTGTGTCTTAGGTTTCCGACAGCTTGAACGAGCTTATCAAGAGGTT 4801
Db 4758 GUUUGUUAACAUAUUGUGCUUAUAC---ACACAGUUGAGACGCGUUAUGGAGGU 4814
Qy 4802 CATAAAAACCGGATTTGATGTTGTTGTTGTTTAAATGTTGTTAAACAATTTTGTGAT 4861
Db 4815 CAUAAAGCGCCCGCCUCCAGGUUCUGUUAUUAUAAAGUCUGGUGAAGAUUUGUCUGAU 4874
Qy 4862 AAATTTTATTAGAACTTTGTTTAAATGGCTGTAGTCTCAGAGATCTGTCAAAAT 4921
Db 4875 AAAGTUCUUUUAAGAUUUGUUAUGGCGUCUUAAGUUAUAAAGAAAGUUAUUAU 4934
Qy 4922 TAGCGAGTTTATGATCTTTCGAAAACAGATGAGATCTTCCGGCATTCATGACTAAGT 4981
Db 4935 CAUAGAUUUUAUCCAGCUGACAAAUAUGAGAGAAUUAUCCGUGAUUUAUCCUUGU 4994
Qy 4982 CAAGAGTTTATAGATAATCGACTGTGGAACAAGTATTGCTGTAAAGATGATGTTTC 5041
Db 4995 AAAGAGUUAUUGUUAUCCAAAGUUAUAAUUAUUGUUAUGAAGAAAGUAGUUAUUG 5054
Qy 5042 TGATGTAGATTATTAAGGTTTAAAGTTAGTTAAGAAAGGTTATGTTGCTTAGCTGA 5101
Db 5055 AGAGUGAAACCTUCUUAAGAGUUAUAGCUUAUUAUAGUUAUAGCUCUGUUAUAGCGG 5114
Qy 5102 TTTGCTAGTGTCTGGGAGTGAATCTCCCGATACTGCCGTGTGTGTGTGTGCTGAGTTG 5161
Db 5115 UUGUGUCUACGCGCGAGUGGAAACUUGCCUGACAAUUGCAGAGAGAGUGUGAGUGUG 5174
Qy 5162 TATTGTAGATAGAGAAATGAAAAAGAGTAAGGAAAGCAACGCTGGGTGCGTATACGCCCC 5221
Db 5175 UCUGUGGACAAAGGAUAGGAAAGAGCGGAGCGGAGGAGGAGGAGGAGGAGGAGGAGG 5234

GenCore version 5.1.1.6
Copyright (c) 1993 - 2005 CompuGen Ltd.

.OM nucleic - nucleic search, using sw model

Run on: January 17, 2005, 05:56:43 ; Search time 12563.7 Seconds
(without alignments)
18431.958 Million cell updates/sec

Title: US-09-551-494-5
Perfect score: 6355
Sequence: 1 gatgtttaatagttttcga.....taacgcgcgtagcgcccca 6355

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 32822875 seqs, 18219865908 residues

Total number of hits satisfying chosen parameters: 65645750

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : EST:*

1: gb_est1:*

2: gb_est2:*

3: gb_hic:*

4: gb_est3:*

5: gb_est4:*

6: gb_est5:*

7: gb_est6:*

8: gb_gssi:*

9: gb_gss2:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
C 1	303.4	4.8	702	4	BM068217 KS08018F1
C 2	299.8	4.7	681	4	BM068406 KS08023B0
C 3	299.6	4.7	700	4	BM067779 KS08011B0
C 4	299.6	4.7	717	4	BM068367 KS08022E0
C 5	297.6	4.7	668	4	BM068523 KS08025D0
C 6	296.6	4.7	663	4	BM068294 KS08020D0
C 7	296.2	4.7	663	4	BM068316 KS08020F1
C 8	295.8	4.7	701	4	BM067993 KS08015C0
C 9	295.6	4.7	665	4	BM067343 KS08003C0
C 10	293.8	4.6	660	4	BM067946 KS08014D0
C 11	293.6	4.6	693	4	BM068203 KS08018F0
C 12	292.8	4.6	656	4	BM068536 KS08025F0
C 13	292.6	4.6	668	4	BM068286 KS08020C0
C 14	292.6	4.6	677	4	BM067601 KS08007H1
C 15	292.4	4.6	679	4	BM067465 KS08005D1
C 16	291.2	4.6	655	4	BM067937 KS08014C1
C 17	290.2	4.6	654	4	BM067953 KS08014E0
C 18	290.2	4.6	694	4	BM068482 KS08024F0
C 19	289.8	4.6	652	4	BM067532 KS08006H0
C 20	289.8	4.6	665	4	BM068348 KS08022C0
C 21	287.4	4.5	658	4	BM067756 KS08010E0
C 22	287.2	4.5	651	4	BM067815 KS08011H0
C 23	286.2	4.5	667	4	BM068012 KS08015E1
C 24	286	4.5	654	4	BM067516 KS08006E0

25	285.8	4.5	688	4	BM067317	KS08002H0
C 26	285	4.5	644	4	BM067900	KS08013F1
C 27	284.8	4.5	647	4	BM067464	KS08005D0
C 28	284.8	4.5	673	4	BM067541	KS08007A0
C 29	284	4.5	644	4	BM067352	KS08003D0
C 30	283.4	4.5	662	4	BM067303	KS08002F0
C 31	283	4.5	645	4	BM068308	KS08020F0
C 32	282.2	4.4	694	4	BM068447	KS08024A0
C 33	281.8	4.4	645	4	BM067543	KS08007A1
C 34	278.8	4.4	653	4	BM068271	KS08019H1
C 35	278.6	4.4	637	4	BM068485	KS08024G0
C 36	278.6	4.4	650	4	BM068116	KS08017B1
C 37	277.8	4.4	660	4	BM068227	KS08018H0
C 38	276.8	4.4	651	4	BM068498	KS08025A0
C 39	276.6	4.4	689	4	BM067450	KS08005S0
C 40	270.4	4.3	646	4	BM068487	KS08024G0
C 41	270.2	4.3	642	4	BM067359	KS08003E0
C 42	270.2	4.3	655	4	BM067772	KS08011A0
C 43	269	4.2	629	4	BM067413	KS08004D1
C 44	268.2	4.2	625	4	BM067817	KS08011H1
C 45	268.2	4.2	638	4	BM067339	KS08003B1

ALIGNMENTS

RESULT 1
BM068217/c 702 bp mRNA linear EST 11-SEP-2002
LOCUS KS08018F11 KS08 Capsicum annuum cDNA, mRNA sequence.
DEFINITION BM068217
ACCESSION BM068217.1 GI:22788322
VERSION EST.
KEYWORDS
SOURCE Capsicum annuum
ORGANISM Capsicum annuum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Capsicum.

REFERENCE 1 (bases 1 to 702)
AUTHORS Lee,S., Kim,S.-Y., Chung,Y.-H., Shin,H.-J., Goh,S.-H., Pai,H.-S., Hur,C.-G. and Choi,D.
TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum annuum L.) and Sequence Analysis in Relation to Hypersensitive Response Against Pathogen
JOURNAL Unpublished (2001)
COMMENT Contact: Doil Choi
Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yusong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doil@mail.kribb.re.kr
High quality sequence stop: 702.

FEATURES
source
1..702
/organism="Capsicum annuum"
/mol_type="mRNA"
/cultivar="Hang Keun"
/db_xref="taxon:4072"
/tissue_type="anther"
/dev_stages="10 weeks after germination"
/clone_lib="KS08"
/note="Vector: pBluescript SK(-)"

ORIGIN
Query Match 4.8%; Score 303.4; DB 4; Length 702;
Best Local Similarity 65.7%; Pred. No. 2.9e-63;
Matches 442; Conservative 0; Mismatches 231; Indels 0; Gaps 0;

OY 3146 AAGGGCTATAAGAAATGTGAACACCGTTTCATGAGATCCAAAGGAGAAACCTTTGAAGATGTG 3205
DB 679 AGGGGTACGAGAGATGATACATCCGTTTCATGAGGTGCAAGGGGAAACCTTTGAACACGTC 620

```
QY 3206 TCGCTGGTCAGATTGACGCAATCCACTGACTCTGATTCCAAAGTCTTCCCGCATGTT 3265
Db      |||
619 TCATTAGTGAGACTGACCGCTACACCGTGGGAATAATTTCAAAGCAGAGTCCGACCTG 560
QY 3266 CTAGTCGCTCTGACTAGACACAAAGAGCTTCAATATTAACCGTAGTGTAGATCCT 3325
Db      |||
559 TTGGTCTCGTGTCTAGGAGATACAAAGGTCAATCAATATTAACACTGTTGTACTAGATGCA 500
QY 3326 TTAGTACAGATAAATAGTGTCTCTTTAAAGCTCTCTTAAAGCTCTCTTTAGAAATGATATG 3385
Db      |||
499 GTCGTTTCAAGTCTTAGAGATTGGAGTGTGAGTAGTACTGTTAGATATGATACAAA 440
QY 3386 GTAGAAAGCAGGTAGTAGATAGCAATTAACAGATGAGTGTGAGTGTCTTAAAGGTCAATATCTC 3445
Db      |||
439 GTTGATGTCTGACTCAATAGCAATTAACAGATAGAAATCGGTGTACAAAGGTGTTAACTT 380
QY 3446 TTGTGGCAACACCTAAATCAGGAGACTTTCCAGATCTACAGTCTTATACAGATGATGTC 3505
Db      |||
379 TTCGTGCGAGCCCCGAAACAGAGATGTTCTGACATGCAATATTAATGACAAAGTGT 320
QY 3506 CTCCTGTGTAATAGTACTATACTTAACAGATGATGCTGTTPACCATGAGGTACGTGAT 3565
Db      |||
319 TTGCGCGGAAACAGTACTATACTTAATGAGTATGATGCTGTAACTATGCAATACGAGAA 260
QY 3566 AATAGTCTTAATGTGAGGATGTGTTCTTGAATTTTCCAAAGATTAATCCGATGCCAAG 3625
Db      |||
259 AATAATTTGAATGTCAAGGATGTGTTGGATATGTCGAATGTCGAGTGCCTCTCCGAGA 200
QY 3626 GAGGTGAACCATGTCTAGAGCCAGTTTGGCTACCGCGGGAACCGCCAGGCTGCA 3685
Db      |||
199 GAATCTGAGACAGATTGAACTGTGATCAGGACTGCTGTGTAAGAAACCTCGAAACCT 140
QY 3686 GGACTACTCGAAATCTGGTTCGAATGATTAAGAAATTTCAACGCCAGACCTGACG 3745
Db      |||
139 GGATTTGGAAACTTGGTCGCGATGATCAAAAGAAATTTCACTCTCCGAAATTAATA 80
QY 3746 GGGAGATGACATTTAGAGACCGCATCTGTTGTAGTAGATAAGTTTTTTGTAGCTAT 3805
Db      |||
79 GGGGTGCTGTGACATCGAAGACACCGCTTCTCTAGTAGTAGATAAGTTTTTTGTAGCTATC 20
QY 3806 TTTATTAATAAAG 3818
Db      |||
19 TTTATTAATAAAG 7
```

```
RESULT 2
BM068406/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Capsicum.
1 (bases 1 to 681)
Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S.,
Hur, C.-G. and Choi, D.
Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
annuum L.) and Sequence Analysis in Relation to Hypersensitive
Response Against Pathogen
Unpublished (2001)
Contact: Doil Choi
Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doil@mail.kribb.re.kr
High quality sequence stop: 681.
Location/Qualifiers
```

FEATURES

source

```
1. 681
/organism="Capsicum annuum"
/mol_type="mRNA"
/cultivar="Hang Keun"
/db_xref="taxon:4072"
/tissue_type="anther"
/dev_stage="10 weeks after germination"
/clone_lib="KS08"
/note="Vector: pBluescript SK(-)"
```

ORIGIN

```
Query Match      4.7%; Score 299.8; DB 4; Length 681;
Best Local Similarity 65.4%; Pred. No. 2.2e-62;
Matches 439; Conservative 0; Mismatches 232; Indels 0; Gaps 0;
QY 3146 AAGGCTATAGAATGTGAACACCGTTTCATGAGATCCAAAGGAGAAACCTTTGAAGATGTG 3205
Db      |||
678 AGGGGTTTACGAAATGTGCATACCTGTTTCATGAGGTGCAAGGGGAAACGTTTGAAGACGTC 619
QY 3206 TCGCTGGTCAGATTGACGGCAACTCCACTGACTCTGATTTTCCAAGTCTTCCCGCATGTT 3265
Db      |||
618 TCATTAGTGAGACTGACCGCTTACACCGTGGGAATAATTTCAAAGCAGAGTCCGACCTG 559
QY 3266 CTAGTCGCTCTGACTAGACACAAAGAGCTTCAATATTAACACCGTAGTGTAGATCCT 3325
Db      |||
558 TTGCTCTCGTGTCTAGGCATACAAAGTCAATCAATATTAACACTGTTGTACTAGATGCA 499
QY 3326 TTAGTACAGATAAATAGTGTGTTCTTCTTTAAGTCTCTTCTTTAGAAATGATATG 3385
Db      |||
498 GTCGTTTCAAGTCTTAGAGATTGGAGTGTGTAGTAGTACCTGTTAGATATGATACAA 439
QY 3386 GTAGAAGCAGGTAGTAGATAGCAATTAACAGATGATGATGTTTCAAAGGTCAATACTC 3445
Db      |||
438 GTTGATGTGCGACTCAATAGCAATTAACAGATAGATCGGTGTACAAAGGTGTTAACTT 379
QY 3446 TTGTGCGAACACCTAAATCAGGAGACTTTCCAGATCTACAGTCTTATTACGATGTATGC 3505
Db      |||
378 TTCGTCGACGCCCGAAACAGGAGATGTTTCTGACATGCAATATTAATGATGACAAAGTGT 319
QY 3506 CTCCTCGTAAATAGTACTACTTAACAAGTATGCTGTACCATGAGTGTACGTTACGTTAT 3565
Db      |||
318 TTGCGGGGAAACAGTACTACTTAAAGTATGATGCTGTAACTGCAAAATACAGAGAA 259
QY 3566 AATAGTCTTAATGTGAAGGATGTGTTCTTGAATTTTCCAAAGATATTCGATGCCCAAAG 3625
Db      |||
258 AATAATTTGAATGTCAAGGATGTGTTGATATGTCGAAGTCGCTCTTCCGAGA 199
QY 3626 GAGGTGAACCATGTCTAGAGCCAGTTTTCGTTACCGCGGGAACCGCCAGGCTGCA 3685
Db      |||
198 GAATCTGAGACGACATTTGAACACCTGTGATCAGGACTGCTGCTGAAACACCTCGAAAACT 139
QY 3686 GGACTACTCGAAATCTGGTTCGAATGATTAAGAAATTTCAACGCCAGACCTGACG 3745
Db      |||
138 GGATTTGGAAACTTGGTCGCGATGATCAAAAGAAATTTCACTCTCCGAAATTAATA 79
QY 3746 GGGAGATGACATTTAGAGACCGCATCTGTTGTAGTAGATAAGTTTTTTGTAGCTAT 3805
Db      |||
78 GGGGTGCTGCACATCGAAGACACCGCTTCTCTAGTAGTAGATAAGTTTTTTGTAGCTATC 19
QY 3806 TTTATTAATAA 3816
Db      |||
18 TTTATTAATAA 8
```

RESULT 3

```
BM057779/c
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Capsicum.
1 (bases 1 to 681)
Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S.,
Hur, C.-G. and Choi, D.
Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
annuum L.) and Sequence Analysis in Relation to Hypersensitive
Response Against Pathogen
Unpublished (2001)
Contact: Doil Choi
Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doil@mail.kribb.re.kr
High quality sequence stop: 681.
Location/Qualifiers
```

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Capsicum.

REFERENCE 1 (bases 1 to 700)

AUTHORS Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S., Hur, C.-G. and Choi, D.

TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum annuum L.) and Sequence Analysis in Relation to Hypersensitive Response Against Pathogen

JOURNAL Unpublished (2001)

COMMENT Contact: Doil Choi

Genome Research Center and National Center for Genome Information

Korea Research Institute of Bioscience and Biotechnology

P.O. Box 115, Yusong, Taejeon, 305-600, Republic of Korea

Tel: 82-42-860-4340

Fax: 82-42-860-4309

Email: doil@mail.kribb.re.kr

High quality sequence stop: 700.

FEATURES

source

1..700

/organism="Capsicum annuum"

/mol_type="mRNA"

/cultivar="Hang Keun"

/db_xref="taxon:4072"

/tissue_type="anther"

/dev_stage="10 weeks after germination"

/clone_lib="KS08"

/note="Vector: pBluescript SK(-)"

ORIGIN

Query Match 4.7%; Score 299.6; DB 4; Length 700;
Best Local Similarity 64.6%; Pred. No. 2.5e-62;
Matches 446; Conservative 0; Mismatches 244; Indels 0; Gaps 0;

QY 3110 ACTTTCACTCAGCGTGATAAATTTGAGTTAGAGGAGAGGGCTATAAGATGCTGAACACC 3169
DB 699 ACATTCACTCAATCAGCAAGTCATTGCTGCTCAAGGGGTTACAGATGTCATACC 640

QY 3170 GTTCATGATCAAGGAGAAACCTTTGAAGATGTGTCGTGTCAGATTTGACGGCACT 3229
DB 639 GTTCATGAGTCAAGGGGAAACGTTTGAAGCGTCTCATTTAGTGAAGTCAAGCGCTACA 580

QY 3230 CCACTGACTCTGATTTCCAGTCTCCCGCATGTTCTAGTGCCTCTGACTAGACACACA 3289
DB 579 CCGGTGGGAATAATTTCAAGCAGAGTCGCGACCTGTTGCTGCTCGTTGCTAGGCATACA 520

QY 3290 AAGAGCTTCAATATTTACACCGTAGTGTGTAGATCCTTTAGTACAGATTAATGATGTTG 3349
DB 519 AGGTCAATCAATATTTACACTGTTGCTAGATGAGTGTTCAGTGTCTTAGAGATTG 460

QY 3350 TCTTCTTTAAGCTCTCTCTTTTGAATAATGTATGTAGTGAAGCAGGTAGTAGACAA 3409
DB 459 GAGTGTGTGAGTGTACCTGTTAGATATGTACAAAGTTGATGTGCTCAATAGCAA 400

QY 3410 TTACAGATGGATCGAGTGTCTAAAGTCAATATCTTTTGTGGCAACACCTAAATCAGGA 3469
DB 399 TTACAGATGAATCGGTGTACAAAGTGTAACTTTTCGTGCGAGCCCCGAAACAGGA 340

QY 3470 GACTTCCAGATCACTAGTCTTATACGATGTATGCTCCCTGTTAATAGTACTACTT 3529
DB 339 GATGTTTCTGACATGCAATATTACTATGACAAGTGTTCGCGGAAACAGTACTACTT 280

QY 3530 AACAAATGATGCTGTGTACCATGAGGTTACGTGATATAGTCTTAATGTAAGGATTGT 3589
DB 279 AATGATGATGCTGTGTAATGCAATATGCAATACGAGAAATATTTGAAATGCAAGTGT 220

QY 3590 GTTCTTGATTTTCCAAAGATATTCGATGCCAAAGAGGTGAACCATGCTAGAGCCA 3649
DB 219 GTGTGGATATGTGGAATCGGTGCTCTTCCGAGAGATCTGAGACGACATTGAAACCT 160

QY 3650 GTTTTCGTCACCGCGCAACCGCAAGGCTGACGAGTACTCGAAATCTGTTGCA 3709
DB 159 GTGATCAGGACTGCTGCTGAAAAACCTCGAAACCTGGATTGTTGGAAACCTTGGTCGG 100

QY 3710 ATGATTAAGAAATTTTCAACGCCACGACCTGACGGGACGATTTGACATTTAGAGACCC 3769
DB 99 ATGATCAAGAAATTTTCAACTCTCCGAAATTAATAGGGTCTGTTGACATCGAAGACCC 40

QY 3770 GCATCTGTTGTAGTAGATAAGTTTTTCAT 3799

DB 39 GCTTCTCTAGTAGTAGATAAGTTTTTCAT 10

RESULT 4

BM068367/c

LOCUS

DEFINITION

ACCESSION

VERSION

KEYWORDS

SOURCE

ORGANISM

Capsicum annuum

Capsicum annuum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;

Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;

asterids; lamids; Solanales; Solanaceae; Capsicum.

REFERENCE 1 (bases 1 to 717)

AUTHORS Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S., Hur, C.-G. and Choi, D.

TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum annuum L.) and Sequence Analysis in Relation to Hypersensitive Response Against Pathogen

JOURNAL Unpublished (2001)

COMMENT Contact: Doil Choi

Genome Research Center and National Center for Genome Information

Korea Research Institute of Bioscience and Biotechnology

P.O. Box 115, Yusong, Taejeon, 305-600, Republic of Korea

Tel: 82-42-860-4340

Fax: 82-42-860-4309

Email: doil@mail.kribb.re.kr

High quality sequence stop: 717.

FEATURES

source

1..717

/organism="Capsicum annuum"

/mol_type="mRNA"

/cultivar="Hang Keun"

/db_xref="taxon:4072"

/tissue_type="anther"

/dev_stage="10 weeks after germination"

/clone_lib="KS08"

/note="Vector: pBluescript SK(-)"

ORIGIN

Query Match 4.7%; Score 299.6; DB 4; Length 717;
Best Local Similarity 64.8%; Pred. No. 2.5e-62;
Matches 460; Conservative 0; Mismatches 249; Indels 1; Gaps 1;

QY 3108 TAACTTTCACCTCAGCTGATAAATTTGAGTTAGAGGAGGCTATAAGATGTGAACA 3167
DB 717 TTACATTCACCTCAATCAGACAGTCATTTGCTCTCAAGGGGTTACGAAGTGTGATA 658

QY 3168 CCGTTCATGAGATCCAAAGGAGAAACCTTTGAAGATGTGTGCTGTGCTAGATTGACGGCAA 3227
DB 657 CCGTTCATGAGGTGCNAGGGGAAAGCTTTGAAGAGCTCTCATTTAGAGACTGACGCTTA 598

QY 3228 TCCCACTGACTCTGATTTCCAAAGCTTCCCGCATGTTCTAGTCTGCTCTGATAGACACA 3287
DB 597 CACCGTGGGAATAATTTCAAAGCAGAGTCGCGACCTGTGTGCTCTGTTGTCTAGGCATA 538

QY 3288 CAAAGAGCTTCAATATTTACACCGTAGTGTAGATCCCTTTAGTACAGATAATTAGTGATT 3347
DB 537 CAAAGTCAATCAATATTTACACTGTGTACTAGATGCAAGTCCGTTTTCAGTGTAGAGATT 478

QY 3348 TGTCTTCTTTAAGCTCTCTCTTTTGAATAATGTATGTAGAAAGCAGAGTGTAGTAGTAGC 3407
DB 477 TGGAGTGTGTAGTAGTTACCTGTTTAGATATGTACAAAGTTGATGTGTGCTCAATAGC 418

QY 3408 AATTACAGATGATCGAGTGTCAAGG-TCATAATCTCTTTGTGGCAACACCTTAATCA 3466
|||||
Db 417 AATTACAGATAGATCGGTGTACAAAGGNTGTTAACTTTTCGTCGACGCCGGAACA 358
|||||
QY 3467 GGAGACTTCCAGATCTACAGTCTTATTACAGTGTATGCCCTCCCTGGTAATAGTACTATA 3526
|||||
Db 357 GGAGATGTTTCTGACATGCAATATTACTATGACAAAGTGTTCGCCGGGAACAGTACTATA 298
|||||
QY 3527 CTTAAACAGTATGATGCTGTTACCATGAGGTTTACGTGATAATAGTCTTAATGTGAAGGAT 3586
|||||
Db 297 CTTATAGTATGATGCTGTAACTATGCAATATACGAGAAATATTGAAATGCAAGGAT 238
|||||
QY 3587 TGTGTTCTGTGATTTTCCAAAGTATCCGATGCCAAAGAGGTGAACACCATGTAGAG 3646
|||||
Db 237 TGTGTTGATATGTCGAAGTCGGTGCCTCTTCCGAGAGAATCTGAGACGACATTGAA 178
|||||
QY 3647 CCAGTTTTCGTACCGCGCGGACCCGCAAGGCTGCAGGACTACTCGAANAATCTGTT 3706
|||||
Db 177 CCGTGATCAGGACTGCTGCTGAAACCTCGAACAACCTGGAATGTTGGAANAATCTGTC 118
|||||
QY 3707 GCAATGATTAAAGAAATTTCAACGACACGACACCTGACGGGACGATTGACATTGAGAGC 3766
|||||
Db 117 GCGATGATCAAGAAATTTCACTCTCCGAAATTAATAGGGTGTGATGACATCGAGAC 58
|||||
QY 3767 ACCGATCTGTGTAGTAGATAAGTTTTCATAGCTATTTTATTAATAA 3816
|||||
Db 57 ACCGCTTCTCTAGTAGTAGATAAGTTTTCATAGCTATTTTATTAATAA 8
|||||

RESULT 5
BM068523/c 668 bp mRNA linear EST 11-SEP-2002
LOCUS KS08025D04 KS08 Capsicum annum cDNA, mRNA sequence.

DEFINITION BM068523
ACCESSION BM068523
VERSION BM068523.1 GI:22788643

KEYWORDS EST.
SOURCE Capsicum annum
ORGANISM Capsicum annum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Capsicum.

AUTHORS 1 (bases 1 to 668)
Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S.,
Hur, C.-G. and Choi, D.

TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
annuum L.) and Sequence Analysis in Relation to Hypersensitive

JOURNAL Response Against Pathogen
COMMENT Unpublished (2001)
Contact: Doil Choi

Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea

FEATURES
source
1. 668
/organism="Capsicum annum"
/mol_type="mRNA"
/cultivar="Hang Keun"
/db_xref="taxon:4072"
/tissue type="anther"
/dev stages="10 weeks after germination"
/clone lib="KS08"
/note="Vector: pBluescript SK(-)"

ORIGIN
Query Match 4.7%; Score 297.6; DB 4; Length 668;
Best Local Similarity 65.9%; Pred. No. 7.7e-62;
Matches 432; Conservative 0; Mismatches 224; Indels 0; Gaps 0;

QY 3159 ATGTGAACACCGTTTCATGAGATCCAAAGAGAAACCTTTGAAGATGTGCTGTCAGAT 3218

Db 663 ATGTGCAATACCGTTTCATGAGTGCAGGGGAACGTTTGAAGACGTTCTCATTAGTGAGAC 604
|||||
QY 3219 TGACGGCAATCCCACTGACTCTGATTTTCAAGTCTTCCCGCATGTTCTAGTCGCTCTGA 3278
|||||
Db 603 TGACGCCCTACACCCGTTGGGAATATTTCAAGCAGAGTCCGACCTGTTGCTCTGTTGT 544
|||||
QY 3279 CTAGACACACAAAGAGCTTCAAAATATTACACCGTAGTGTTTAGATCCTTTAGTACAGATAA 3338
|||||
Db 543 CTAGGCATACAAAGTCAATCAAAATATTACACTGTTGACTAGATGCACTGTTTTCAGTGC 484
|||||
QY 3339 TTAGTGATTTGCTCTTTTAAGCTCCTTCTTTTAGAAATGATATGTTAGAACGAGTA 3398
|||||
Db 483 TTAGAGATTTGGAGTGTGTGAGTAGTTTACCTGTTTAGATATGTACAAAGTGTGATGTGCGA 424
|||||
QY 3399 GTAGATAGCAATACAGATGCGATGTTTCAAGGTTCATAATCTCTTTCTGTGCAACAC 3458
|||||
Db 423 CTCAATAGCAATACAGATAGATCGGTGTACAAAGGTGTTAACCTTTTCGTCGAGGCC 364
|||||
QY 3459 CTAAATCAGGAGACTTTCAGATCTACAGTCTTATTACGATGTATGCTCCCTCGTATAA 3518
|||||
Db 363 CGAAACAGGAGATGTTTCTGACATGCAATATTACTATGACAAGTGTTCGCGGGAACA 304
|||||
QY 3519 GTACTATCTTAAACAGTATGATGCTGTTTACCATGAGGTTACGTTGATATAGTCTTTATG 3578
|||||
Db 303 GTACTATCTTAAATGAGTATGATGCTGTAATATGCAAAATACGAGAAATTAATTTGAATG 244
|||||
QY 3579 TGAAGGATGTTGTTCTGATTTTCCAAAGATTTTCGATGTCGCAAGAGGTGAACCAT 3638
|||||
Db 243 TCAAGGATGTTGTTGATGATGTCGAGTGGTGTCTCTTCGAGAGATCTCGAGGGA 184
|||||
QY 3639 GTCTAGAGCCAGTTTTTCGCTACCGCGGGAACCGCAAGGGCTGCAGGACTACTCGAAA 3698
|||||
Db 183 CATTGAAACCTGTGATCAGGACTGCTGCTGAAACCTCGAACAACCTGGATTGTTGAAA 124
|||||
QY 3699 ATCTGGTTGCAATGATTAAGAAATTTCAACGCCACGACCTGACGGGAGGATTGACA 3758
|||||
Db 123 ACTTGGTGGCGATGATCAAAAGAAATTTCAACTCTCCGAAATTAATAGGGGTCTGTGACA 64
|||||
QY 3759 TTGAGACACCGCATCTGTTGTAGTAGTAAGTTTTCATAGTCTTTTATTTATTTAAA 3814
|||||
Db 63 TCGAAGACACCGCTTCTCTAGTAGTAGTAAGTTTTCATAGTCTTTTATTTAAA 8
|||||

RESULT 6
BM068294/c 663 bp mRNA linear EST 11-SEP-2002
LOCUS KS08020D04 KS08 Capsicum annum cDNA, mRNA sequence.

DEFINITION BM068294
ACCESSION BM068294
VERSION BM068294.1 GI:22788414

KEYWORDS EST.
SOURCE Capsicum annum
ORGANISM Capsicum annum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamids; Solanales; Solanaceae; Capsicum.

AUTHORS 1 (bases 1 to 663)
Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S.,
Hur, C.-G. and Choi, D.

TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
annuum L.) and Sequence Analysis in Relation to Hypersensitive

JOURNAL Response Against Pathogen
COMMENT Unpublished (2001)
Contact: Doil Choi

Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea

FEATURES
source
1. 663
/organism="Capsicum annum"
/mol_type="mRNA"
/cultivar="Hang Keun"
/db_xref="taxon:4072"
/tissue type="anther"
/dev stages="10 weeks after germination"
/clone lib="KS08"
/note="Vector: pBluescript SK(-)"

```

/organism="Capsicum annuum"
/mol_type="mRNA"
/cultivar="Hang Keun"
/db_xref="taxon:4072"
/tissue_type="anther"
/dev_stages="10 weeks after germination"
/clone_lib="KS08"
/note="Vector: pBluescript SK(-)"

ORIGIN
Query Match      4.7%; Score 296.6; DB 4; Length 663;
Best Local Similarity 65.8%; Pred. No. 1.4e-61;
Matches 431; Conservative 0; Mismatches 224; Indels 0; Gaps 0;

QY 3163 GAACACCGTTCATGAGATCCAAAGGAGAAACCTTTGAAGATGTCGCTGGTCAGATTGAC 3222
DB 663 GCATACCGTTCATGAGTGCAGAGGAGAAACGTTTGAAGACGCTCTCATTAGTGAGACTGAC 604

QY 3223 GGCACACTCCACTGACTCTGATTTCCAGTCTTCCCGCATGTTCTAGTCGCTCTGACTAG 3282
DB 603 GCTACACCGTGGGAATTAATTCAAAGCAGAGTCGCGACCTGTTGGTCTCGTTCCTAG 544

QY 3283 ACACACAAGAGCTTCAATATATTACACCGTAGTGTGTAGATCTTTAGTACAGATAATTAG 3342
DB 543 GCATACAGGTCAATCAATATATTACCTGTTGTACTAGTGCAGTCTTTCAGTCTTAG 484

QY 3343 TGAATTTGCTTCTTTAAGCTCTTCTTTTGAAGATGATATGTTAGAGCAGGTTAGTAG 3402
DB 483 AGATTTGGAGTGTGTGAGTAGTACCTGTTAGATATGTACAAAGTTGATGTGCTGACTCA 424

QY 3403 ATAGCAATTTACAGATGATGTCGATCTTCAAGGTCTAATCTCTTTGTTGGCAACCTTAA 3462
DB 423 ATAGCAATTTACAGATGATGTCGATCTTCAAGGTCTAATCTCTTTGTTGGCAACCTTAA 364

QY 3463 ATCAGGAGACTTTCAGATCTACAGTCTTATTAACGATGATGCTTCCCTGTTGTAATGATAC 3522
DB 363 AACAGAGATGTTCTGACATCAATATTAATCTATGACAGTGTTCGCGGAGAACAGTAC 304

QY 3523 TATACCTTAAAGATGATGTCGTTTACATGAGGTTACGTGATATATAGTCTTAAATGTGAA 3582
DB 303 TATACCTTAAAGATGATGTCGTTTACATGAGGTTACGTGATATATAGTCTTAAATGTGAA 244

QY 3583 GGAATGTGTTCTGATTTTCCAAAGATTTCCGATGTCGATGTCGATGTCGATGTCGATGTC 3642
DB 243 GGAATGTGTTCTGATTTTCCAAAGATTTCCGATGTCGATGTCGATGTCGATGTCGATGTC 184

QY 3643 AGAGCCAGTTTTCGCTACCGCGCGGAAACCGCAAGGGCTGCAGGACTACTCGAAATCT 3702
DB 183 GAAACCTGTGATCAGACTGCTGCTGAAACCTCGAAACCTGGAATGTTGGAAACTT 124

QY 3703 GGTTCGAATGATTAAGAAATTTCAACGACAGACTGACGGGAGGATGACATGGA 3762
DB 123 GGTTCGAATGATTAAGAAATTTCAACGACAGACTGACGGGAGGATGACATGGA 64

QY 3763 GAGCACCCACTGTTGTAGTAGATAGTATTTTTCATAGCTATTTTATTAATAAAA 3817
DB 63 AGACACCCGCTCTCTAGTAGTAGATAGTATTTTTCATAGCTATTTTATTAATAAAA 9

RESULT 7
BM068316/c BM068316 663 bp mRNA linear EST 11-SEP-2002
LOCUS KS08020F12 KS08 Capsicum annuum cDNA, mRNA sequence.
DEFINITION BM068316
ACCESSION BM068316
VERSION BM068316.1 GI:22788436
KEYWORDS EST.
SOURCE Capsicum annuum
ORGANISM Capsicum annuum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Capsicum.
REFERENCE 1 (bases 1 to 663)
AUTHORS Lee,S., Kim,S.-Y., Chung,Y.-H., Shin,H.-J., Goh,S.-H., Pai,H.-S.,

```

```

Hur,C.-G. and Choi,D.
Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
annuum L.) and Sequence Analysis in Relation to Hypersensitive
Response Against Pathogen
Unpublished (2001)
Contact: Doil Choi
Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doil@mail.kribb.re.kr
High quality sequence stop: 663.
Location/Qualifiers
1..663
/organism="Capsicum annuum"
/mol_type="mRNA"
/cultivar="Hang Keun"
/db_xref="taxon:4072"
/tissue_type="anther"
/dev_stages="10 weeks after germination"
/clone_lib="KS08"
/note="Vector: pBluescript SK(-)"

ORIGIN
Query Match      4.7%; Score 296.2; DB 4; Length 663;
Best Local Similarity 65.8%; Pred. No. 1.7e-61;
Matches 430; Conservative 0; Mismatches 223; Indels 0; Gaps 0;

QY 3165 ACACCGTTCATGAGATCCAAAGGAGAAACCTTTGAAGATGTCGCTGGTCAGATTGACGG 3224
DB 661 ATACCGTTCATGAGTGCAGGAGGAAACGTTTGAAGACGCTCTCATTAGTGAGACTGACGC 602

QY 3225 CAATCCACTGACTCTGATTTCCAAAGTCTTCCCGCATGTTCTAGTCCCTCTGACTAGAC 3284
DB 601 CTACACCGTGGGAATTAATTTCAAAGCAGAGTCCGACCTGTTGCTCTGTTGTCTAGGC 542

QY 3285 ACACAAAGAGCTTCAATATTTACACCGTAGTGTGTAGATCTTTAGTACAGATAATTAGTG 3344
DB 541 ATACAAAGGTCAATCAATATTTACACTGTTGTACTAGATGCAGTCTGTTTCAGTGTCTAGAG 482

QY 3345 ATTTGCTCTTTTAAGCTCTTCTTTTGAAGATGTTATATGTAAGCAGGTAGTAGAT 3404
DB 481 ATTTGGAGTGTGTGAGTAGTACCTGTTAGATATGTAACAAAGTGTGATGTGCTGACTCAAT 422

QY 3405 AGCAATTAACAGATGATGTCAGTGTTCAAAGGTCTAATCTCTTTGTGCAACACCTAAAT 3464
DB 421 AGCAATTAACAGATGATGTCAGTGTTCAAAGGTGTTAACTTTTCGTCGAGCCCGGAAAA 362

QY 3465 CAGGAGACTTTCAGATCTACAGTCTTATTAACGATGATGTCCTCCCTGGTAAATAGTACTA 3524
DB 361 CAGGAGATGTTTCTGACATGCAATATTAATGTAACAGTGTTCGCGGAAACAGTACTA 302

QY 3525 TACTTAAACAGTATGATGCTGTTACCATGAGTTACGTGATAATAGTCTTAAATGTGAAGG 3584
DB 301 TACTTAAAGATGATGCTGTTAACTATGCAATACAGAAATTAATTTGATGTCAAGG 242

QY 3585 ATTGTGTTCTTGAATTTTCCAAAAAGTATTCGATGTCGCAAGAGGTGAACCATCTCTAG 3644
DB 241 ATTGTGTTGGATATGTCGAAAGTCGGTGCCTCTCTCGAGAGATCTCGAGACGACATTGA 182

QY 3645 AGCCAGTTTTCGTTACCGCGCGGAAACCGCAAGGGCTGCAGGACTACTCGAAATCTGG 3704
DB 181 AACCTGTGATCAGGACTGCTCTGTAAGAAACCTCGAAAAACCTGGATTTGTTGAAAAA 122

QY 3705 TTGCAATGATTAAGAAATTTCAACGACACAGACTCTGACGGGAGGATGACATTGAGA 3764
DB 121 TCGCATGATCAAAAGAAATTTCAACTCTCCGAATTAATAGGGGTCTGTGACATCGAAG 62

QY 3765 GCACCGCATCTGTTGTAGTAGATAAGTTTTTGTAGTCTATTTTATTAATAAAA 3817
DB 61 ACACCGCTCTCTAGTAGTAGATAGTATTTTGTGATGATGCTATTTTATTAATAAAA 9

```



```

RESULT 8
BM067993/c
LOCUS      BM067993      701 bp      mRNA      linear      EST 11-SEP-2002
DEFINITION KS08015C03 KS08 Capsicum annuum cDNA, mRNA sequence.
ACCESSION  BM067993
VERSION     BM067993.1  GI:22788113
KEYWORDS   EST.
SOURCE     Capsicum annuum
ORGANISM   Capsicum annuum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamids; Solanales; Solanaceae; Capsicum.
REFERENCE  1 (bases 1 to 701)
AUTHORS   Lee,S., Kim,S.-Y., Chung,Y.-H., Shin,H.-J., Goh,S.-H., Pai,H.-S.,
            Hur,C.-G. and Choi,D.
TITLE     Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
            annuum L.) and Sequence Analysis in Relation to Hypersensitive
            Response Against Pathogen
JOURNAL   Unpublished (2001)
COMMENT   Contact: Doil Choi
            Genome Research Center and National Center for Genome Information
            Korea Research Institute of Bioscience and Biotechnology
            P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
            Tel: 82-42-860-4340
            Fax: 82-42-860-4309
            Email: doil@mail.kribb.re.kr
            High quality sequence stop: 701.
FEATURES   source
            Location/Qualifiers
                1..701
                /organism="Capsicum annuum"
                /mol_type="mRNA"
                /cultivar="Hang Keun"
                /db_xref="taxon:4072"
                /tissue_type="anther"
                /dev_stage="10 weeks after germination"
                /clone_lib="KS08"
                /note="Vector: pBluescript SK(-)"
ORIGIN
Query Match      4.7%; Score 295.8; DB 4; Length 701;
Best Local Similarity 65.2%; Pred. No. 2.2e-61;
Matches 435; Conservative 0; Mismatches 23; Indels 0; Gaps 0;
QY 3146 AAGGCTATAAAGTGTGAACACCGCTTCATGAGATCCAAAGGAGAAACCTTTGAAGATGTG 3205
DB 674 AGGGGTTACGAAGAGTACATACCGTTTCATGAGGTGCAAGGGGAAACGTTTGAAGACGTC 615
QY 3206 TCGCTGGTCAAGTTCACGCGCACTCCACTGACTCTGATTTCCAAGTCTTCCCGCATGTT 3265
DB 614 TCATTAGTGAGACTGACGCGCTACACCGCTGGGAATAATTTCAAAGCAGAGTCCGCAACCTG 555
QY 3266 CTAGTCGCTCTGACTAGACACAAAGAGTTTCAAATATTACCGTAGTGTAGATCCT 3325
DB 554 TTGCTCTGTTGTCTAGGCATACAGGTCAATCAATTAATTAATCTGTTGTAAGTACGA 495
QY 3326 TTAGTACAGATAAATPAGTGTGTTGCTTTTAAAGCTCCTTCCCTTTTGAAGAATGTATATG 3385
DB 494 GTCGTTTCAGTGCTTAGAGATTGGAGTGTGTGAGTAGTTACCTGTTAGATATGTACAAA 435
QY 3386 GTAGAACAGGTAGTAGATAGCAATACAGATGATGATGAGTTCAAAGGTCAATATCTC 3445
DB 434 GTTGAATGTGCTGACTCAATAGCAATACAGATGATGATGATGATGATGATGATGATGATG 375
QY 3446 TTGTGGCAACACCTTAATCAGGAGACTTCCAGATCTACAGTCTTATTACGATGTATGC 3505
DB 374 TTGTCGCGAGCCCGGAAACAGGAGATGTTTCTGACATGCAATATTACTATGACAGTGT 315
QY 3506 CTCCTGTGTATAGTACTATATCTTAACAAGTATGATGCTGTTTACATGAGTTACGTGAT 3565
DB 314 TTGCGGGGAAACAGTACTATATCTTAATGAGTATGATGCTGTTAACTATGCAATACGAGAA 255
QY 3566 AATAGTCTTAATGTGAAGGATTGTTCTTCTGATTTTCCAAAGATTTCCGATGCCAAG 3625

```

```

Db 254 AATAATTGAATGTCAAGGATTGTGTGTGATATATCGAAGTCGGTGCTCTTCCGAGA 195
QY 3626 GAGGTGAACCATCTAGAGCCAGTTTTCGTACCGGGCGGAACCGCCAAAGGCTGCA 3685
DB 194 GAATCTGAGACGACATTGAAACCTGTATCAGGACTGCTCTGAAAAACCTTCGAAAACT 135
QY 3686 GGACTACTCGAAAAATCTGGTTGCAATGATTAAAAAGAAATTTCAACGACACGACCTGACG 3745
DB 134 GGATGTTGGAAAACTTGGTCGCGATGATCAAGAAGAAATTTCAACTCTCCGAAATTAATA 75
QY 3746 GGGACGATTGACATGAGAGACCGCATCTGTTCTAGTAGATAAGTTTTTTTGTAGCTAT 3805
DB 74 GGGTCGTTGACATCGAAGACACCGCTTCTCTAGTAGTAGATAAGTTTTTTTGTAGCTAT 15
QY 3806 TTTTATTA 3812
DB 14 TTTTATTA 8

```

```

RESULT 9
BM067343/c
LOCUS      BM067343      665 bp      mRNA      linear      EST 11-SEP-2002
DEFINITION KS08003C07 KS08 Capsicum annuum cDNA, mRNA sequence.
ACCESSION  BM067343
VERSION     BM067343.1  GI:22787463
KEYWORDS   EST.
SOURCE     Capsicum annuum
ORGANISM   Capsicum annuum
            Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
            Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
            asterids; lamids; Solanales; Solanaceae; Capsicum.
REFERENCE  1 (bases 1 to 665)
AUTHORS   Lee,S., Kim,S.-Y., Chung,Y.-H., Shin,H.-J., Goh,S.-H., Pai,H.-S.,
            Hur,C.-G. and Choi,D.
TITLE     Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
            annuum L.) and Sequence Analysis in Relation to Hypersensitive
            Response Against Pathogen
JOURNAL   Unpublished (2001)
COMMENT   Contact: Doil Choi
            Genome Research Center and National Center for Genome Information
            Korea Research Institute of Bioscience and Biotechnology
            P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
            Tel: 82-42-860-4340
            Fax: 82-42-860-4309
            Email: doil@mail.kribb.re.kr
            High quality sequence stop: 665.
FEATURES   source
            Location/Qualifiers
                1..665
                /organism="Capsicum annuum"
                /mol_type="mRNA"
                /cultivar="Hang Keun"
                /db_xref="taxon:4072"
                /tissue_type="anther"
                /dev_stage="10 weeks after germination"
                /clone_lib="KS08"
                /note="Vector: pBluescript SK(-)"
ORIGIN
Query Match      4.7%; Score 295.6; DB 4; Length 665;
Best Local Similarity 65.7%; Pred. No. 2.4e-61;
Matches 430; Conservative 0; Mismatches 224; Indels 0; Gaps 0;
QY 3161 GTGAACACCGCTTCATGAGATCCAAAGGAGAAACCTTTGAAGATGTGTCGCTGAGATG 3220
DB 665 GTGCATACCGTTTCATGAGGTGCAAGGGGAAACGTTTGAAGACGCTCTCATAGTACTG 606
QY 3221 ACGCCAATCCACTGACTCTGATTTTCAAGTCTTCCCGCATGTTCTAGTCGCTGACT 3280
DB 605 ACGCCTACACCGCTGGGAATAATTTCAAAGCAGAGTCCGACACCTGTGTGCTGCTGCT 546
QY 3281 AGACACAAAGAGCTTCAATATATTACCGTAGTGTAGATCCTTTAGTACAGATAATT 3340

```

Db 545 AGGCATCAAGTCAATCAATATATACACTGTTGTACTAGATGCAGTCTGTTTCAGTGTCT 486
QY 3341 AGTAGTTGTCCTTTTAAGCTCCTTCTTTAGAAATGATATATGATAGAGCAGGTAGT 3400
Db 485 AGAGATTTGGAGTGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGTAGTGT 426
QY 3401 AGATAGCAATATACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 3460
Db 425 CAATAGCAATATACAGATGATGATGATGATGATGATGATGATGATGATGATGATG 366
QY 3461 AAATCAGGAGATCTTCCAGATCTACAGATCTTATACGATGATGATGATGATGATG 3520
Db 365 AAAACAGGAGATGTTCTGACATGCAATATATCTATGATGATGATGATGATGATG 306
QY 3521 ACTATCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3580
Db 305 ACTATCTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 246
QY 3581 AAGGATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3640
Db 245 AAGGATGTTGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 186
QY 3641 CTAGAGCCAGTCTTCCGATGATGATGATGATGATGATGATGATGATGATGATGATG 3700
Db 185 TTGAACCTGATCAGGACTGCTGCTGAAACCTCGAAACCTGATGTTGGAAC 126
QY 3701 CTGTTGCAATGATTAAGAAATTTCAACGACGATGATGATGATGATGATGATGATG 3760
Db 125 TTGTTGCGGATGATCAAAAGAAATTTCAACTCTCCGAAATTAAGGGGTGCTTGACATC 66
QY 3761 GAGACACCGCATGTTGTAGTATGATGATGATGATGATGATGATGATGATGATGATG 3814
Db 65 GAAGACACCGCTTCTCTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 12

RESULT 10
BM067946/c
LOCUS KS08014D08 KS08 Capsicum annuum cDNA, mRNA linear EST 11-SEP-2002
DEFINITION BM067946
ACCESSION BM067946
VERSION BM067946.1 GI:22788066
KEYWORDS EST.
SOURCE Capsicum annuum
ORGANISM Capsicum annuum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Capsicum.
Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S.,
Hur, C.-G. and Choi, D.
Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
annuum L.) and Sequence Analysis in Relation to Hypersensitive
Response Against Pathogen
Unpublished (2001)
Contact: Doil Choi
Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doll@mail.kribb.re.kr
High quality sequence stop: 660.
Location/Qualifiers
1. .660
/organism="Capsicum annuum"
/mol_type="mRNA"
/cultivar="Hang Keun"
/db_xref="taxon:4072"
/tissue_type="anther"
/dev_stage="10 weeks after germination"
/clone_lib="KS08"
/note="Vector: pBluescript SK(-)"

FEATURES
source

ORIGIN

Query Match 4.6%; Score 293.8; DB 4; Length 660;
Best Local Similarity 65.3%; Pred. No. 6.7e-61;
Matches 427; Conservative 0; Mismatches 227; Indels 0; Gaps 0;
QY 3165 ACACGGTTTCATGAGATCCAAAGGAGAAACCTTTTGAAGATGTCTCGTGTGATGAGCGG 3224
Db 660 ATACCGTTTCATGAGTGCAGGGGAAAGCTTTTGAAGAGCTCTCATTTAGTGAAGTAAAGC 601
QY 3225 CAATCCACTGACTCTGATTTTCCAAAGTCTTCCCGCATGTTTCTAGTCTGCTCTGACTAGAC 3284
Db 600 CTACACCGTGGGAATAATTTTCAAAGCAGAGTCCGACCTGTGCTCTGCTGTCTAGGC 541
QY 3285 ACACAAAGAGCTTCAATATTTACACCGTAGTGTGTAGATCCCTTTAGTACAGATAATTAGT 3344
Db 540 ATCAAGGTCAATCAATATTTACACTGTTGTACTAGATGCAGTCTGTTTCAGTGTCTAGAG 481
QY 3345 ATTTGTCTTCTTTAAGCTCTCTCTTTTAAAGATGTATATGTTAGAAATGTTATGTTAGAACGAGTACTAGAT 3404
Db 480 ATNTGGAGTGTGTAGTAGTTTACCTGTTAGATATGTACANAGTTGATGTGCTGACTCAAT 421
QY 3405 AGCAATTAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3464
Db 420 AGCAATTAACAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 361
QY 3465 CAGGAGATTTCCAGATCTACAGTCTTATTAAGATGATGATGATGATGATGATGATGATG 3524
Db 360 CAGGAGATGTTTCTGACATGCAATATTTACTATGATGCAAGTGTGTCGCGGAAACAGTACTA 301
QY 3525 TACTTAAACAGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 3584
Db 300 TACTTAAAGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 241
QY 3585 ATTTGTTTCTTGTATTTTCCAAAGATGATTTCCGATGATGATGATGATGATGATGATG 3644
Db 240 ATTTGTTTCTTGTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 181
QY 3645 AGCCAGTTTTCGTTACCGCGGGAACCGCAAGGCTGCAGGACTACTCGAAATCTGG 3704
Db 180 AACCTGTGATCAGGACTCTCTGTAAGAAACCTCGAAACCTGGAATGTTGTGGAACCTGG 121
QY 3705 TTGCAATGATTAAAGAAATTTTCAACGCCAGACACTGACGGGAGCAATTTGACATTTGAGA 3764
Db 120 TCGGATGATCAAAAGAAATTTTCAACTCTCCGATTAATAGGGGTCTGTGACATCGAAG 61
QY 3765 GCACCGCATCTGTTGTAGTAGATAAGTTTTTGTAGTAGTATTTTATTTAAAAAG 3818
Db 60 ACACCGCTTCTCTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAGTAG 7

RESULT 11
BM068203/c
LOCUS KS08018E01 KS08 Capsicum annuum cDNA, mRNA linear EST 11-SEP-2002
DEFINITION BM068203
ACCESSION BM068203
VERSION BM068203.1 GI:22788308
KEYWORDS EST.
SOURCE Capsicum annuum
ORGANISM Capsicum annuum
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Capsicum.
Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S.,
Hur, C.-G. and Choi, D.
Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
annuum L.) and Sequence Analysis in Relation to Hypersensitive
Response Against Pathogen
Unpublished (2001)
Contact: Doil Choi
Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea

Tel: 82-42-860-4340
 Fax: 82-42-860-4309
 Email: doilemail.kr@kribb.re.kr
 High quality sequence stop: 693.

FEATURES

source
 1..693
 /organism="Capsicum annuum"
 /mol_type="mRNA"
 /cultivar="Hang Keun"
 /db_xref="taxon:4072"
 /tissue_type="anther"
 /dev_stage="10 weeks after germination"
 /clone_lib="KS08"
 /note="Vector: pBluescript SK(-)"

ORIGIN

Query Match 4.6%; Score 293.6; DB 4; Length 693;
 Best Local Similarity 64.3%; Pred. No. 7.5e-61;
 Matches 440; Conservative 0; Mismatches 244; Indels 0; Gaps 0;

QY 3113 TTCACTCAGGCTGATAAATTTCAGTTAGAGGAGAGGCTATAAGAAATGTGAACACCGTT 3172
 DB |||||
 QY 692 TTCACTCAATCAGACAGATCATTTCTCTCAAGGGGTACCAAGATGTGCTACTGTT 633
 DB |||||
 QY 3173 CATGAGATCCAAAGGAAACCTTTGAAGATGTTCGCTGGTCAAGTCAACGCAACTCCA 3232
 DB |||||
 DB 632 CATGAGGTGCAAGGGGAAACGTTTGAAGACGCTCTCATTAGTGAGACTGACGCTACACCC 573
 QY 3233 CTGACTCTGATTTCCAAAGCTTCCCGCATGTTCTAGTCGCTCTGACTAGACACAAAG 3292
 DB |||||
 DB 572 GTGGGAATATTTCAAGACAGAGTCGCGACCTGTTGGTCTCGTGTCTAGGCATACAAAG 513
 QY 3293 AGCTTTCAATATTACACCGTAGTGTAGATCCCTTTAGTACAGATAAATTAGTATTGCT 3352
 DB |||||
 DB 512 TCAATCAATATTACACTGTTGTACTAGATGACGTCGTTTCAGTCTTAGAGATTGAG 453
 QY 3353 TCTTTAAGCTCTCTCTTTGAAGATGTATGTAGTAGAGAGGAGTAGATAGCAATTA 3412
 DB |||||
 DB 452 TGTGTGAGTAGTACTCTGTATGATATGACAAAGTTGATGTCTGACTCAATAGCAATTA 393
 QY 3413 CAGATGGATGCAAGTGTCAAGAGTCATAATCTTTTGGGCAACACCTTAATCAGGAC 3472
 DB |||||
 DB 392 CAGATAGATCGGTGTACAAGGTGTTAACTTTTCGCGAGCCCGGAAACAGGAGAT 333
 QY 3473 TTTCAGATCTACAGTCTTATPACGATGTATGCCCTCCCTGGTAACTAGTACTTAAAC 3532
 DB |||||
 DB 332 GTTCTGACATGCAATATTACTATGACAAAGTGTTCGCGGAAACAGTACTATCTTAAT 273
 QY 3533 AAGTATGATGCTTTACCATGAGTTAGTGATATAGTCTTAATGTGAAGATTGTT 3592
 DB |||||
 DB 272 GAGTATGATGCTGTAACTATGCAATACGAGAAATATATTTGAATGTCAAGGATTGTGTG 213
 QY 3593 CTGTGATTTTCCAAAGTATTCCGATGCCAAGAGGAGTGAACCATGCTAGAGCCAGTT 3652
 DB |||||
 DB 212 TTGATATGTGCAAGTCGGTGCCTCTTCGAGAGATCTGAGACGATTTGAACCTGTG 153
 QY 3653 TTGCGTACCGGGGGAACCGCAAGGCTCGAGGCTACTCGAAATCTGGTTGCAATG 3712
 DB |||||
 DB 152 ATCAGGACTGCTGTGAAGAAACCTCGAAACCTCGGATTGTTGGAAACCTTGGTCGCGATG 93
 QY 3713 ATTTAAAGAAATTTCAACGCAACCGACCTGACGGGAGATTGACATTTGAGACACCGCA 3772
 DB |||||
 DB 92 ATCAAGAAATTTTCACTCTCCGCAATTAATAGGGGTCGTTGACATCGAAGACACCGCT 33
 QY 3773 TCTGTGTAGTAGATAAGTTTTTT 3796
 DB |||||
 DB 32 TCTCTAGTAGTAGATAAGTTTTTT 9

RESULT 12
 BM068536/c
 LOCUS
 DEFINITION KS08025F01 KS08 Capsicum annuum cDNA, mRNA EST 11-SEP-2002

ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM

BM068536
 BM068536.1 GI:22788656
 EST.
 Capsicum annuum
 Capsicum annuum

REFERENCE

1 (bases 1 to 656)
 Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S.,
 Hur, C.-G. and Choi, D.

AUTHORS

Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
 annuum L.) and Sequence Analysis in Relation to Hypersensitive
 Response Against Pathogen

TITLE

Unpublished (2001)
 Contact: Doll Choi

JOURNAL

Genome Research Center and National Center for Genome Information
 Korea Research Institute of Bioscience and Biotechnology
 P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea

COMMENT

Tel: 82-42-860-4340
 Fax: 82-42-860-4309
 Email: doilemail.kr@kribb.re.kr
 High quality sequence stop: 656.

FEATURES

source
 1..656
 /organism="Capsicum annuum"
 /mol_type="mRNA"
 /cultivar="Hang Keun"
 /db_xref="taxon:4072"
 /tissue_type="anther"
 /dev_stage="10 weeks after germination"
 /clone_lib="KS08"
 /note="Vector: pBluescript SK(-)"

ORIGIN

Query Match 4.6%; Score 292.8; DB 4; Length 656;
 Best Local Similarity 65.7%; Pred. No. 1.2e-60;
 Matches 426; Conservative 0; Mismatches 222; Indels 0; Gaps 0;

QY 3170 GTTCATGAGATCCAAAGGAGAAACCTTTGAAGATGTGTCGCTGTGTCAGATTGACGCAACT 3229
 DB |||||
 DB 656 GTTCATGAGGTGCAAGGGGAAACGTTTGAAGACGTTCTATTAGTAGACTGACGCTTACA 597
 QY 3230 CCACCTGACTCTGATTTTCCAAAGTCTTCCCGCATGTTCTAGTCGCTCTGACTAGACACACA 3289
 DB |||||
 DB 596 CCGTGGGGAATATTTCAAGACAGAGTCGCGACCTGTTGCTCTCGTTGTCTAGCATACA 537
 QY 3290 AAGAGCTTTCAAAATATTACACCGTAGTGTAGATCCTTTTAGTACAGATAAATTAGTATTG 3349
 DB |||||
 DB 536 AGGTCAATCAAAATATTACACTGTTTGTACTAGATGCACTCGTTTCAGTGTCTTAGAGATTG 477
 QY 3350 TCTTCTTTAAGCTCCTTCCCTTTTAGAAATGTATATGTTAGAACGAGTAGTAGTAGCAAA 3409
 DB |||||
 DB 476 GAGTGTGTGAGTAGTTACCTGTTAGATATGTACAAGATTGATGTGCGACTCAATAGCAA 417
 QY 3410 TTACAGATGGATCAGTGTTCAAAAGTCAATAATCTCTTTGTGGCAACACCTTAATCAGGA 3469
 DB |||||
 DB 416 TTACAGATAGAATCGGTGTACAAAGGTGTTAACTTTTCGTCGAGCCCGGAAACAGGA 357
 QY 3470 GACTTTCCAGATCTACAGTTCATTAGCATGTATGCTCCCTCGTAACTAGTACTATACCTT 3529
 DB |||||
 DB 356 GATGTTTCTGACATGCAATATTACTATGACAAAGTGTTCGCGGAAACAGTACTATACCTT 297
 QY 3530 AACAAAGTATGATGCTGTACCATGAGTTACGTTACGTTAATAAGTCTTAAATGTGAAGGATTGT 3589
 DB |||||
 DB 296 AATGATATGATGCTGTACTACTATGCAAAATACGAGAAATAATTTGAATGTCAAGGATTGT 237
 QY 3590 GTTCTGTGATTTTCCAAAAGTATTCGATGCGCAAGAGGTGAAACCATGTCTTAGAGCCA 3649
 DB |||||
 DB 236 GTGTGTGATATGTGGAAGTCGGTGCCTCTTCCGAGAGAAATCTGAGACGACATTTGAAACCT 177
 QY 3650 GTTTTGGGTACCGCGGGAACCGCAAGGCTGCGAGGACTACTCTGAAATCTGTTTGA 3709
 DB |||||

Db 176 GTGATCAGGACTGCTGCTGAAAAACCTCGAAAAACCTGGATTGTTGGAAAACTTGTGTCGG 117

Qy 3710 ATGATTAAGAAATTTCAACGCCACACAGCTGACGGGAGGAGTATGACATGAGAGCAC 3769

Db 116 ATGATCAAGAAATTTCAACTCTCCGAAATTAATAGGGGTGTTGACATCGAAGACACC 57

Qy 3770 GCATCTGTTGTAGTAGATAGTATTTTTCATAGCTATTTATTAATAAAA 3817

Db 56 GCTTCTCTAGTAGTAGTAAGTATTTTGTATGATCATACTTTATTAATAAAA 9

RESULT 13

BM068286/c

LOCUS BM068286 668 bp mRNA linear EST 11-SEP-2002

DEFINITION KS08020C01 KS08 Capsicum annuum cDNA, mRNA sequence.

ACCESSION BM068286

VERSION BM068286.1 GI:22789406

KEYWORDS EST.

SOURCE Capsicum annuum

ORGANISM Capsicum annuum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Capsicum.

AUTHORS Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S., Hur, C.-G. and Choi, D.

TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum annuum L.) and Sequence Analysis in Relation to Hypersensitive Response Against Pathogen

JOURNAL Unpublished (2001)

COMMENT Contact: Doil Choi
Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doil@mail.kribb.re.kr
High quality sequence stop: 668.
Location/Qualifiers

FEATURES

source 1..668

/organism="Capsicum annuum"

/mol_type="mRNA"

/cultivar="Hang Keun"

/db_xref="taxon:4072"

/tissue_type="anther"

/dev_stage="10 weeks after germination"

/clone_lib="KS08"

/note="Vector: pBluescript SK(-)"

ORIGIN

Query Match 4.6%; Score 292.6; DB 4; Length 668;

Best Local Similarity 65.6%; Pred. No. 1.3e-60;

Matches 427; Conservative 0; Mismatches 224; Indels 0; Gaps 0;

Qy 3160 TGTGAACCGTTTATGATGATCAAGGAGAAACCTTTGAAGATGTCGCTGGTCAGATT 3219

Db 659 TGTGATACCGTTTATGATGATCAAGGAGAAACCTTTGAAGATGTCGCTGGTCAGACT 600

Qy 3220 GACGCACTCCACTGACTCTGATTTCCAAAGTCTCCCGCATGTTAGTCGCTGAC 3279

Db 599 GACGCTACACCCGGGGAATAATTTCAAAGCAGAGTCGCGACCTGTTGGTCTCGTTGTC 540

Qy 3280 TAGACACAAAGAGCTTCAATATTTACACCGTAGTGTAGATCTTTAGTACAGATAAT 3339

Db 539 TAGGCATACAGGTCAATCAATATTTACACTGTTGATGATGAGTCGTTTCACTGCT 480

Qy 3340 TAGTGAATTTGCTTCTTTAAGCTCTCTCTTTTAGAAATGTATATGGTAGAGCAGGTAG 3399

Db 479 TAGAGATTTGGAGTGTGTGAGTAGTACCTGTTAGATATGTACAAAGTTGATGTCGAC 420

Qy 3400 TAGATAGCAATTAAGATGAGTGTGTTCAAAGTCAATATCTCTTTTGGGAAACACC 3459

Db 419 TCAATAGCAATTAAGATGAGTGTGTTCAAAGTGTGTTACCTTTTTCGTCGACGCC 360

Qy 3460 TAAATCAGGAGACTTTCAGATCTACAGTTCTATATAGATGTATGCCCTCCCTGTAATAG 3519

Db 359 GAAAACAGGAGATGTTTCTGACATGCAATATTTACTATGACAAGTGTTTGCCGGGAAACAG 300

Qy 3520 TACTATATCTTAAACAGTATGATGCTGTGTACCATGAGGTTAGTGATATATGCTTTAATGT 3579

Db 299 TACTATATCTTAAACAGTATGATGCTGTGTACCATGAGGTTAGTGATATATGCTTTAATGT 240

Qy 3580 GAAGGATTTGTTCTTGTATTTTCCAAAAGTATTCGATGCCAAAGGAGGTGAACACCATG 3639

Db 239 CAAGGATTTGTTGTTGGATATGTGCAAGTCGGTGCCTCTCCGAGAGAATCTGAGACGAC 180

Qy 3640 TCTAGAGCCAGTTTGTGTACCGGGCGGACCCGCAAGGGCTGCAGACTACTCTGAAAA 3699

Db 179 ATTGAACACCTGTGATCAGGACTGCTGCTGAAAAACCTCGAAAAACCTGATGTTGGA 120

Qy 3700 TCTGTTGCAATGATTAAGAAAAATTTCAACGCGACACGACCTGACGGGGAGCATTTGACAT 3759

Db 119 CTTGTCGCGATGATCAAGAAAAATTTCACTCTCCGAAATTAATAGGGGTCGTTGACAT 60

Qy 3760 TGAGAGCACCCGATCTGTTGTAGTAGATAAGTATTTTGTATAGCTATTTTAT 3810

Db 59 CGAAGACACCCGCTCTCTAGTAGTAGATAAGTATTTTGTATGATCATACTTTAT 9

RESULT 14

BM067601/c

LOCUS BM067601 677 bp mRNA linear EST 11-SEP-2002

DEFINITION KS08007H10 KS08 Capsicum annuum cDNA, mRNA sequence.

ACCESSION BM067601

VERSION BM067601.1 GI:22787721

KEYWORDS EST.

SOURCE Capsicum annuum

ORGANISM Capsicum annuum

REFERENCE Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; asterids; lamids; Solanales; Solanaceae; Capsicum.

AUTHORS Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S., Hur, C.-G. and Choi, D.

TITLE Generation of Expressed Sequence Tags from Hot Pepper (Capsicum annuum L.) and Sequence Analysis in Relation to Hypersensitive Response Against Pathogen

JOURNAL Unpublished (2001)

COMMENT Contact: Doil Choi
Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doil@mail.kribb.re.kr
High quality sequence stop: 677.
Location/Qualifiers

FEATURES

source 1..677

/organism="Capsicum annuum"

/mol_type="mRNA"

/cultivar="Hang Keun"

/db_xref="taxon:4072"

/tissue_type="anther"

/dev_stage="10 weeks after germination"

/clone_lib="KS08"

/note="Vector: pBluescript SK(-)"

ORIGIN

Query Match 4.6%; Score 292.6; DB 4; Length 677;

Best Local Similarity 65.3%; Pred. No. 1.3e-60;

Matches 430; Conservative 0; Mismatches 229; Indels 0; Gaps 0;

Qy 3146 AAGCGCTATAGATGTGAACACCGTTTCATGATGCCAAGGAGAAACCTTTGAAGATGTG 3205

Db 668 AGGGTTACGAAGATGTGCATACCTGTTTATGAGGTGCAAGGGGAAACGTTTGAAGACGTC 609

```
QY 3206 TCGCTGCTCAGATTGACGGCACTCCACTGACTCTGATTTCAGAGTCTTCCCGCAGATTT 3265
Db |||||
QY 608 TCATTAGTGAGACTGACCGCTACACCCGTGGGAATAATTTCAAAGCAGAGTCCGCACCTG 549
Db |||||
QY 3266 CTAGTCGCTCTGACTAGACACACAAGAGCTTCAATATTAACCGTAGTGTAGATCCT 3325
Db |||||
QY 548 TTGGTCTCGTTGCTTAGGCATACAAAGGTCATCAATATTAACCTGTTGTAAGTACGA 489
QY 3326 TTAGTACAGATAAATTAGTGATTGTTCTTTTAAGCTCTTCTTTTGAAGATGTAATG 3385
Db |||||
QY 488 GTCGTTTCAGTGCTTAGAGATTGGAGTGTGAGTAGTCTTACCTGTTAGATATGACAAA 429
QY 3386 GTAGAAGCAGGTAGTAGACAAATTAAGATGAGTGTTCAAAGGTCATTAATCTC 3445
Db |||||
QY 428 GTTGATGTGTCGACTCAATAGCAATTAAGATAGAAATCGGTGTACAAAGGTTTAAACCTT 369
QY 3446 TTTGTGGCAACACCTTAATCAGGAGCTTCCAGATCTACAGTCTTAATTAAGATGTGTC 3505
Db |||||
QY 368 TTGCTGCGAGCCCGGAAACAGGAGATGTTTCTGACATGCAATATTAATGACAAAGTGT 309
QY 3506 CTCCCTGCTAATAGTACTTAACTTAAAGATGATGCTGTACCATGAGGTTACGTGAT 3565
Db |||||
QY 308 TTGCGGGAACAGTACTTACTTAACTGATGATGCTGTAACTATGCAATACAGAA 249
QY 3566 AATAGCTTAATGGAAGATGTTCTTGAATTTTCCAAAGATTTCCGATGCCAAAG 3625
Db |||||
QY 248 AATAATTTGAATGTCAAGGATTTGTGTTGATATGTCGAAGTCGGTCCCTCTTCCGAGA 189
QY 3626 GAGGTGAACCATGCTAGAGCCAGTTTGGTACCGCGGCGGACCCGAGGGCTGCA 3685
Db |||||
QY 188 GAATCTGAGACGACATTTGAACCTGTGATCGGAGTCTGCTGTGAAACACCTCGAAACCT 129
QY 3686 GGACTACTCGAAATCTCGTTGCAATGATTAAGAAATTTCAACGACACGAGCTGACG 3745
Db |||||
QY 128 GGATTTGGAACCTTGTGTCGATGATCAAGAAATTTCACTCTCCGAAATTAATA 69
QY 3746 GGGAGATGATGACATGAGAGACCGCATCTGTTGATGATAGTAAGTTTTTTGATAGCTA 3804
Db |||||
QY 68 GGGGTCGTTGACATCGAAGACACCGCTTCTCTAGTAGTAGTAAGTTTTTTGATGCATA 10
```

RESULT 15
BM067465/c
LOCUS KS08005D10 KS08 Capsicum annum cDNA, mRNA linear EST 11-SEP-2002
DEFINITION BM067465
ACCESSION BM067465
VERSION BM067465.1 GI:22787585
KEYWORDS EST.
SOURCE Capsicum annum
ORGANISM Capsicum annum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
asterids; lamiales; Solanales; Solanaceae; Capsicum.
1 (bases 1 to 679)
Lee, S., Kim, S.-Y., Chung, Y.-H., Shin, H.-J., Goh, S.-H., Pai, H.-S.,
Hur, C.-G. and Choi, D.

Generation of Expressed Sequence Tags from Hot Pepper (Capsicum
annuum L.) and Sequence Analysis in Relation to Hypersensitive
Response Against Pathogen
Unpublished (2001)
Contact: Doil Choi

Genome Research Center and National Center for Genome Information
Korea Research Institute of Bioscience and Biotechnology
P.O. Box 115, Yuseong, Taejeon, 305-600, Republic of Korea
Tel: 82-42-860-4340
Fax: 82-42-860-4309
Email: doil@mail.kribb.re.kr

High quality sequence stop: 679.
Location/Qualifiers
1..679

FEATURES
source
/organism="Capsicum annum"
/mol_type="mRNA"
/cultivar="Hang Keun"

/db_xref="taxon:4072"
/tissue_type="anther"
/dev_stage="10 weeks after germination"
/clone_lib="KS08"
/note="Vector: pBluescript SK(-)"

ORIGIN

```
Query Match 4.6%; Score 292.4; DB 4; Length 679;  
Best Local Similarity 65.4%; Pred. No. 1.5e-60;  
Matches 428; Conservative 0; Mismatches 226; Indels 0; Gaps 0;  
QY 3146 AAGGGCTATAGATGTCACACCGTTTCATGAGTCCAAAGGAGAACTTTGAAGATGTG 3205  
Db |||||  
QY 664 AGGGGTTACGAAGATGTGCATACCGTTTCATGAGTGCAAGGGGAAACGTTTGAAGACGTC 605  
Db |||||  
QY 3206 TCGCTGCTCAGATTGACGGCAACTCCACTGACTCTGATTTTCCAAAGTCTTCCCCCATGTT 3265  
Db |||||  
QY 604 TCATTAGTGAGACTGACGCTTACACCGTGGGAATATTTCAAAGCAGAGTCCGCACCTG 545  
Db |||||  
QY 3266 CTAGTCGCTCTGACTAGACACACAAGAGCTTCAAAATATTACACGCTAGTGTAGATCCT 3325  
Db |||||  
QY 544 TTGGTCTCGTTGCTAGGCATACAAAGTCAATCAATATTAACCTGTTGTAAGTACGA 485  
Db |||||  
QY 3326 TTAGTACAGATAAATTAGTGTGTTCTTTTAAGCTCCTTCCCTTTTGAAGATGTAATG 3385  
Db |||||  
QY 484 GTCGTTTTCAGTGTCTAGAGATTTGGAGTGTGTGAGTAGTTTACCTGTTAGATATGACAAA 425  
Db |||||  
QY 3386 GTAGAAGCAGGTAGTAGATGACAAATTAACAGATGATGAGTGTTCAAAGGTCATAATCTC 3445  
Db |||||  
QY 424 GTTGATGTGTCGACTCAATAGCAATTAACAGATGAAATCGGTGTACAAAGGTTTAACTT 365  
Db |||||  
QY 3446 TTTGTGCAACACCTAAATCAGGAGCTTTCCAGATCTACAGTCTTATTAAGTGTATGC 3505  
Db |||||  
QY 364 TTTGTCGAGCCCGGAAACAGGAGATGTTTCTGACATGCAATATTAATGACAAAGTGT 305  
Db |||||  
QY 3506 CTCCTCGTAAATAGTACTATATCTTAAACAGTATGATGCTGTACCAGTGGTTCAGTGAT 3565  
Db |||||  
QY 304 TTGCGGGGAAACAGTACTATCTTAAATGAGTATGATGCTGTAACTATGCAAAATACGAGAA 245  
Db |||||  
QY 3566 AATAGTCTTAATGTGAAGGATTTGTTCTTCTGATTTTCCAAAGTATTTCCGATGCCAAAG 3625  
Db |||||  
QY 244 AATTAATTTGAATGTCAAGGATTTGTTGTTGATATGTCGAAGTCGGTGCCTCTTCCGAGA 185  
Db |||||  
QY 3626 GAGGTGAAACCATGCTTAGAGCCAGTGTTCGGTACCGGGGGAACCCGAGGGCTGCA 3685  
Db |||||  
QY 184 GAATCTGAGACGACATTTGAAACCTGTGATCAGGACTGCTGCTGAAAAACCTCGAAAACT 125  
Db |||||  
QY 3686 GGACTACTCGAAATCTGTTGCAATGATTAAGAAATTTCAACGACACGAGCTGACG 3745  
Db |||||  
QY 124 GGATTTGGAAGAACTTTGGTGGCATGATCAAGAAATTTCAACTCTCCGAAATTAATA 65  
Db |||||  
QY 3746 GGGACGATTTGACATTTGAGAGCACCAGCTCTGTTGATGATAGTAAGTTTTTTTGTAT 3799  
Db |||||  
QY 64 GGGGTCGTTGACATCGAAGACACCGCTTCTCTAGTAGTAGTAAGTTTTTTTGTAT 11  
Db |||||
```

Search completed: January 17, 2005, 16:28:44
Job time : 12565.7 secs